

Potential weaknesses in unaccustomed generosity

The large boost of funds for British science is an essential means to arrest its decline. Attention now needs to focus on ensuring these funds are distributed fairly.

There is, as the poet Samuel Butler pointed out, a natural reluctance to look a gift horse squarely in the mouth; but there are occasions when it can be prudent, if only to avoid nasty surprises later on. Such is the case with the extra funds that the British government has announced it is to make available for the support of research in universities and through the research councils over the next three years (see page 307).

Three interrelated aspects of the boost in funds stand out. The first is the central role in which it places the Wellcome Trust, which is contributing £400 million (\$660 million) towards the £1.4 billion of new money announced by the government — including half of the £600 million fund for refurbishing university laboratories and equipment. The second is the de facto power given to the Director General of Research Councils (DGRC), as co-chair of this fund's steering committee, over the way its money is distributed. The third is the heavy emphasis being given, implicitly through Wellcome's involvement in this fund and explicitly in the allocation of new money to the research councils, to the life sciences.

The logic behind that emphasis is clear. The sciences of molecular biology and genomics are among those at which Britain excels, while its pharmaceutical and biotechnology companies have a stronger competitive edge in the market-place than most other sectors. Wellcome's major commitment to genome research is already complemented by the keen interest that those companies have shown in exploiting this research. Economically at least, it makes sense to play to the strongest suit.

But the danger of a continuing decline in the physical sciences —

or at least in those substantial parts of such disciplines that are unlikely to contribute, either directly or indirectly, to the exploitation of the genome — is obvious. To be fair, the scientists at the top of the government's science structure — Sir Robert May, chief scientific adviser, and Sir John Cadogan, the DGRC — have never underestimated the importance of core sciences. Physics and chemistry will no doubt continue to be supported, at least insofar as they underpin economic development and improving the quality of life. A fair implementation of the original case made by Sir John to the Treasury for the science base would be for Wellcome's £300 million to be spent on rebuilding the infrastructure of biomedical sciences, and most of the other £300 million equipment boost to be devoted elsewhere.

Given their influential supporters, chemists and environmental biologists are likely to see their situations improving. But others can anticipate an accustomed and possibly increasing lack of generosity. Although particle physicists have received the long-term backing of the Particle Physics and Astronomy Research Council for involvement in the Large Hadron Collider at CERN, the European Laboratory for Particle Physics, this leaves the council's other clients, mostly astronomers, more exposed to the pressures of government priorities. Meanwhile 'small' physicists — for example, the excellent UK community working on semiconductors and other functional materials — may suffer by comparison with biologists, given the lack of powerful voices from industry or elsewhere speaking up on their behalf. Without such external advocates, the envisaged structure for allocating the new funds, and the all-powerful DGRC in particular, are likely to have difficulty in finding the motivation to give them the necessary support. □

Adult cloning marches on

New results on cloning technology increase the urgency for regulations to ensure its responsible use.

Despite some recent suggestions to the contrary, it now appears that Dolly (see *Nature* 385, 810–813; 1997) was indeed a clone produced from the nucleus of a differentiated cell derived from an adult sheep. Those who questioned that conclusion can now accept it in the light of experiments, including DNA analysis of original tissues, reported in the Scientific Correspondence section of this issue (see page 329). Circumstantial evidence is also added by the fact that the trick has been accomplished again, this time with mice (page 369). The conclusion for developmental biologists remains fascinating: the differentiation of cells is in some cases reversible — the expression of their genomes can be reprogrammed to that of the undifferentiated state. As Davor Solter describes (page 315), the new results sharpen the scientific challenge to be addressed.

But the results also highlight progress in the face of pessimism about the technical obstacles. They throw into sharp relief the response of some researchers to the concerns expressed by others about human cloning, namely that the technological capability is a long way off, and the practical obstacles apparently insuperable. That

response looks like becoming irrelevant all too quickly.

Thus it becomes all the more probable that, where someone is legally allowed to do it, they will. After all, there are, as Solter bullishly describes, potential benefits. And one country where human cloning is in principle legal, despite strong opposition, is the United States. Only federal funding of such work is banned.

Republicans in Congress have attempted to prohibit altogether not only the gestation and bringing to birth of human clones, but also experiments at the pre-implantation stage — an initiative that was successfully derailed by the scientific community. A bill that would permit pre-implantation research has been proposed, but progress is unlikely before Congress breaks up in early October.

Private companies are free to make their move. Debates continue over the protection of the unborn, while the technology moves on. Given the demonstrable potential for moral panic in this field, it is regrettable, and not in the best interests of science, that the world's scientific superpower has so far failed to deliver a satisfactory contract between researchers and its citizens at large on this issue. □



Gore calls for action on climate change as Congress stalls

[WASHINGTON] The Clinton administration is trying to persuade Congress to implement a \$6 billion programme of research and tax credits aimed at curbing the growth in US emissions of greenhouse gases. It hopes that scientific evidence of soaring global surface temperatures, as well as a spate of extreme weather events, will provide ammunition for its efforts.

Vice-president Al Gore last week used a prolonged heatwave and drought in the southern United States, together with flash floods in his home state of Tennessee and forest fires in Florida, as a basis for new attacks on Congress, which is trying to block the programme.

"The evidence of global warming keeps piling up," he told a news conference in Washington on 14 July. "How long is it going to take until people in the Congress get the message?" he added. "People are sweltering out there."

Pausing momentarily to point out that no single weather event can be attributed to global warming, Gore ran through a litany of such events in the United States before unveiling new climate data from the National Oceanic and Atmospheric Administration. These show that the first half of 1998 was the warmest on record, by a considerable margin.

The data indicate that the mean global surface temperature in June exceeded the previous record — set in June 1994 — by 0.4 °F. Global surface temperatures have hit record levels for every month of this year, with the overall mean temperature being 1.3 °F above average for the period 1880–1997 — a deviation from the average that is 50 per cent greater than ever recorded.

Whatever their scientific significance, both the new data and the drought come at a convenient time for Gore. This summer, the administration faces congressional resistance on several fronts to its relatively modest efforts to slow the growth of US carbon emissions.

For example, a clause added to the House version of the Veterans' Affairs and Housing and Urban Development and Independent Agencies bill would prohibit administration officials from "conducting educational outreach or informational seminars on policies underlying the Kyoto Protocol". Gore describes this as a "gag order", and President Bill Clinton is likely to veto any bill containing it.

A more substantial issue is Congress's reluctance to support the Climate Change Technology Initiative. This is the research and development package proposed by the administration in its February budget for



On toast: hungry grasshoppers have descended on Oklahoma's drought-blighted corn crops.

developing technologies to reduce carbon emissions (see *Nature* 391, 619–620; 1998).

The House Appropriations Committee has removed the additional \$450 million allocated under the initiative for next year on research into renewable energy sources and energy efficiency technologies. The Senate would allow about one-third of the spending to proceed.

John Holdren of Harvard University, a

member of the President's Council of Advisors on Science and Technology, and author of a 1997 council report recommending the extra spending on energy research, said he was "appalled" by what he called Congress's "sabotaging" of the research programme.

"It's an absurd argument that the initiative is the premature implementation of Kyoto," Holdren says. "Research and development doesn't commit you to meeting an emissions target."

The fate of the Climate Change Technology Initiative — and of the administration's proposal for \$3.6 billion of tax breaks for energy conservation — will be determined in September, when the budget and tax bills are due to be finalized.

On a more personal level, Gore is almost certain to be the Democrat nominee in the presidential election of 2000, and his long-standing interest in the climate change issue has encouraged strong resistance to action on climate change among the Republicans who control both houses of Congress. "The reality," says one frustrated Democrat staffer on Capitol Hill, "is that anything with Gore's name on it is dead on arrival up here."

Colin Macilwain

Indian atomic chief is refused US visa

[NEW DELHI] Indian scientists are expressing concern that Washington's decision to refuse the entry of Indian atomic and missile scientists into the United States could affect the free exchange of information between the two countries and set back scientific cooperation.

The US decision was made in response to the nuclear tests conducted by India in May. The Indian warnings follow the denial of a US visa to Rajagopalan Chidambaram, chairman of the Atomic Energy Commission (AEC), preventing him from visiting Atlanta to attend the annual meeting of the International Union of Crystallography, of which he is vice-president.

Chidambaram had applied for a visa not as chairman of the AEC but in his individual capacity as a crystallographer. But the US state department has said that the Clinton administration is reviewing its science and technology engagement with both India and Pakistan after the tests, and that the processing of Chidambaram's visa application was "part of this review process".

Officials in Washington say they have no intention of imposing across-the-board visa restrictions on scientists from either

country, and that restrictions will be applied only to individuals engaged in nuclear and missile development.

But Indian scientists fear that the new rules will be extended to scientists in universities or national laboratories whose research may be remotely connected with nuclear or missile development. The US visa policy "will definitely have a nuisance value for us, but we have to accept this as part of life", says Valangiman Ramamurthi, secretary of India's ministry of science and technology.

Chidambaram is not the only top atomic scientist to have been denied a visa by a western country following the Indian nuclear tests. Placid Rodriguez, director of the Indira Gandhi Centre for Atomic Research (IGCAR) at Kalpakkam near Chennai, and his colleague Baldev Raj have both been denied visas to enter Britain. The two researchers had been invited to a Cambridge meeting by a UK publishing company.

In a separate case, Baldev Raj and S. B. Bhoje, director of IGCAR's reactor division, had to abandon planned visits to institutions in Germany after authorities there "suddenly" withdrew their invitations.

K. S. Jayaraman

NIH streamlines research grants process

[WASHINGTON] The leaders of the National Institutes of Health (NIH) have approved a streamlined procedure for research grant applications that will require budgets to be based merely on multiples of \$25,000, not drawn up to the last cent.

The 'modular' grant format was recently given a provisional go-ahead by Harold Varmus, the NIH director, and other institute directors. The final details are expected to be approved by the directors within the next few weeks, with the scheme coming into effect with the round of grants for which applications are due on 1 February 1999.

Diana Jaeger, acting director of NIH's Office of Policy for Extramural Research Administration, says the aim of the streamlined process is to keep the focus of attention "primarily on the science", for applicants and reviewers alike. She says the NIH wants to "disengage" from intense budgetary negotiations that "add very little value to the grants process".

The new format has already been tested by several NIH institutes, and will apply to all investigator-initiated applications requesting less than \$250,000 in annual funding.

Under the scheme, applicants will be asked to provide a total figure for the cost of the project, expressed as a multiple of \$25,000, and a written justification.

The latter includes estimated figures for four categories of expenses: staff, major equipment, laboratory alterations and renovations, and contracts or arrangements as a member of a consortium. Other items, such as travel, supplies and expenses, will no longer have to be itemized.

The National Heart, Lung and Blood Institute (NHLBI) began piloting modular grants in 1995. Ron Geller, the director of its division of extramural affairs, says the process has meant that "the reviewers don't spend a lot of time worrying about whether travel should be cut from \$1,200 to \$800".

He says many investigators "felt overwhelmingly that it saved time, because they didn't have to deal with the detailed dollars in every single category". Geller says his reviewers found the \$50,000 increments used in an initial experiment too blunt.

Will the \$25,000 increments tempt applicants to round their costs upwards? "We'll be looking for that," says Jaeger, adding that the

NIH is ready to cut \$25,000 off any application that seems inflated. "Peer reviewers are also investigators, and they know what it costs to conduct a project." A first-year assessment conducted by the NIH will also watch for trends in average grant costs under the new scheme, she says.

One scientist who won a modular grant under a pilot project welcomes the procedure. Louis Ptacek says he is "very positive" about the process. Ptacek, a geneticist and Howard Hughes Medical Institute investigator at the University of Utah in Salt Lake City, last year won \$200,000 a year for four years from NHLBI and two other institutes to study advanced sleep phase syndrome, a rare sleep disorder.

Ptacek says that it was a "tremendous relief" not to have to contend with the budget detail of a traditional R01 application. "I had to give some general guidelines about how I intended to spend the money, but it didn't ask for the excruciating detail that is expected in an R01." He estimates that the modular grant format could eventually halve the time he spends drawing up figures for his grant applications. **Meredith Wadman**

Bill tightens law against genetic discrimination by health insurers

[WASHINGTON] Republican members of the US Senate are trying to fill the gaps in a 1996 law prohibiting genetic discrimination in health insurance coverage. A bill introduced last week would prohibit health insurers from using genetic information to deny individual insurance policies to people who have not been previously or recently insured.

This group was not protected under the Health Insurance Portability and Accountability Act of 1996, authored by Senator Edward Kennedy (Democrat, Massachusetts) and former Senator Nancy Kassebaum (Republican, Kansas). Importantly, the new bill, the Patients' Bill of Rights Act, also prohibits health insurers — whether for groups or individuals — from using genetic information to set premium costs. It bars insurers from requiring applicants or those already enrolled to take genetic tests or to divulge test results.

The bill, whose supporters are led by Don Nickles (Republican, Oklahoma), gives a broad definition of genetic information that includes not only an individual's genetic tests but also tests taken by that person's family members, as well as family histories of diseases that put the individual at a significantly increased risk of a disease.

The broad definition was applauded by the National Human Genome Research



Nickles: bill would forbid health insurers from using genetic information to set premium costs.

Institute. "We are delighted that family history has been included here," says Kathy Hudson, the institute's assistant director for policy coordination.

But the breadth and reach of the bill's language on genetic discrimination has not gone down well with the Health Insurance Association of America, which says that its members do not currently use genetic information to deny coverage.

Dean Rosen, the association's senior vice-president of policy, says the bill would set a "really terrible" precedent as it would give the federal government authority over the setting of health insurers' rates, an area historically controlled by the states. He also

argues that it would damage the market for individual insurance coverage, which is taken out by about 10 million Americans.

By guaranteeing individual policies to people who have long held back from buying health insurance but then receive genetic test results revealing them to be at increased risk of disease, "it forces people who have... held on to their insurance for years to subsidize the higher health costs of someone who has waited," says Rosen.

The effect of such behaviour in raising premium costs is not of the same concern to insurers in the group market because of its much larger size.

The genetic discrimination provisions are included in a broad bill that aims to protect the rights of patients in an era of managed healthcare. But the bill's other provisions apply only to the 48 million Americans in federally regulated health plans. In contrast, the genetic discrimination language encompasses an additional 100 million Americans in state-regulated plans, as well as individuals not currently insured.

Neither a competing Democratic bill, authored by Senator Tom Daschle (South Dakota) and Congressman John Dingell (Michigan), nor a bill from House Republican leaders, addresses genetic discrimination. It is not clear which of the three bills will prevail. **M. W.**

UK universities get another funding boost

[LONDON] British universities were told last week to expect an extra £300 million (US\$500 million) over the next three years to support research through the block grants they receive from the Department for Education and Employment.

The money comes on top of the two £300 million sums that the government and the Wellcome Trust had earlier announced they are to provide for a new fund to rebuild and refurbish university laboratories over the same period (see *Nature* 394, 209; 1998).

It is also in addition to the extra £400 million for research council projects, much of which is likely to be spent in universities. Although the precise distribution of the funding over the three years has not yet been decided, it represents a significant increase in the amount that universities currently receive directly from the government to support research in their science departments.

There was a broad welcome from the university community last week for the extra money, which, together with the earlier announcement, goes a long way towards meeting suggestions made last year by the Dearing commission on higher education (see *Nature* 388, 413; 1997).

Diana Warwick, chief executive of the Committee of Vice-Chancellors and Principals, for example, described both the new funds for research and a separate commitment by the government to provide an extra 35,000 university places next year as "excellent news".

There was also considerable relief that, at least in principle, the government has decided to retain the 'dual support' system. This is the system under which universities are supposedly left free to decide how to allocate their core government funding to basic scientific infrastructure, with extra money for individual projects being provided through the six research councils on a competitive basis.

"We are delighted that the government is continuing to back the dual support system," says Brian Fender, chair of the Higher Education Funding Council for England, one of the four regional bodies through which government support for universities is distributed, which will jointly receive an extra £50 million for research next year.

But the government has also sent out a clear signal that it intends to reduce the degree of autonomy that universities enjoy in deciding how to spend their research funds. This follows growing concern in government circles that universities do not always spend such funds as it would like (for example, by using them to fund a large number of researchers on short-term contracts, rather than investing in equipment).

Significantly, the Office of Science and Technology (OST), through John Cadogan,



the director general of research councils, will co-chair with Mike Dexter, the director of the Wellcome Trust, the new fund being jointly set up with the trust to cover university equipment and infrastructure.

The Dearing committee had proposed an alternative route in which all research councils would be required to increase the amount of overhead paid automatically on grants to university researchers from its current level of 46 per cent to 60 per cent to cover these types of costs.

But critics of this suggestion argue that, despite the government's continued emphasis on the need for greater 'transparency' in internal accounting procedures, in many British universities these remain insufficiently developed to ensure that the money goes to genuine overhead costs.

The new fund has helped to persuade the Wellcome Trust to shoulder part of the costs of renewing the research infrastructure of British universities — something it had previously been reluctant to do — and has allowed the OST to help target money and keep a close eye on how it is spent on a project-by-project basis.

But it has also raised concerns about how universities are likely to find the money to renew facilities not directly linked to research priorities. "Where will they find the

money for the non-science sections of their libraries, or for sports facilities used by scientists?" asks one observer.

Others point out that last week's statements by the Department for Education lacked any commitment to renew equipment used for teaching, rather than research. Furthermore, the government has done little to reduce the financial disincentives to a research career by, for example, failing to allocate extra money to increase PhD stipends as the research councils had requested (although leaving the councils free to do this if they so choose out of their allocated research budgets).

Behind all this is a concern that the strong emphasis placed by government statements last week on the importance of the life sciences in general — and genomics-related research in particular — could translate in a funding swing away from the physical sciences, with potentially severe long-term consequences for the economy.

Lynne Jones, for example, a Labour member of the House of Commons select committee on science and technology, who trained as a biochemist, reflected such concerns during a parliamentary debate last week when she asked the government for assurances "that physical science will not miss out in the bonanza".

Government officials argue that they are well aware of the potential dangers involved, particularly as the Wellcome Trust is, by virtue of its trust deeds, restricted to funding projects in the biomedical sciences. One response of such officials has been to underline the important contribution made by other disciplines, such as computing and instrumentation engineering, to the success of genome sequencing research.

But many say they will be watching closely and warily how the new money is divided between scientific disciplines, and how what they describe as the excitement of other, physics-based disciplines is reflected in the final resource allocation.

David Dickson

New Zealand enjoys science jamboree

[DUNEDIN] The city of Dunedin in New Zealand's South Island has marked the 150th anniversary of its foundation by Scottish immigrants by mounting the country's largest ever promotion of science. With an attendance of around 25,000, the two-week festival was organized by the local community and museum with the University of Otago, which was established by Scottish settlers in 1869 as the country's first university.

A spectacular finale in the city centre saw 100 people walking unharmed on red hot

charcoal after physicist John Campbell of Canterbury University had given them an explanation of the physics of thermal conductivity.

Expatriate New Zealander Sir Ian Axford, director of the Max Planck Institute for Aeronomy in Germany, commented during the festival on the low proportion of national expenditure New Zealand devotes to research by international standards. He said New Zealanders must increase spending "pretty quick" if they want to become internationally competitive.

Peter Pockley

Plan for government researchers in Japan to work in industry

[TOKYO] Japan's Science and Technology Agency (STA) has asked the government to reduce restrictions on scientists who work at national research institutes as part of its efforts to create a more flexible, open and competitive research environment.

The request, submitted last week to the government's personnel office, asks for 'research bureaucrats' — as researchers at national institutes are called — to be given more freedom to work in different sectors, such as industry and the academic community. At present, such external activities are strictly regulated by civil service law.

The move is in line with the 1996 Basic Law for Science and Technology, which calls for the promotion of collaborative research between industry, government and universities. It also reflects the thrust of last year's administrative reform drive, which called for a substantial shake-up of management styles within nationally run research centres.

The STA also intends to explore the possibility of turning national research institutes into 'independent administrative agencies'.

One of the main reforms being requested by the agency is that researchers be allowed to take lengthy sabbaticals, during which they could take up research different from their normal work, such as projects at profit-making organisations.

The STA hopes that, by reducing bureaucratic obstacles, researchers at national institutes will be able to experience different research environments and improve their standards by carrying out independent research activities.

At the same time, the agency wants the introduction of a system to assess laboratory technicians to allow them to respond to the rapid diversification of scientific fields. The aim is for technicians to be assessed and classified according to their experience and scientific knowledge. The current system merely distinguishes between 'administrative staff' and 'research staff', placing both in one of the lowest salary categories.

The goal is to increase the number of skilled technicians by making it more attractive for them to work in national research institutes. Such a move would be welcomed by scientists, who have long complained about the shortage of skilled technicians.

The STA is also seeking a revision of researchers' salaries to balance them with those of companies and universities, as part of an effort to attract talented researchers.

Scientists welcome the STA's requests. A survey by the agency last year showed that 67 per cent of researchers at national laboratories wanted more freedom. **Asako Saegusa**

Fraud squad files report to prosecutor in Inserm case

[PARIS] Legal authorities in France may launch shortly a more far reaching investigation into a controversial affair involving alleged scientific misconduct than the limited inquiry promised by the Ministry of National Education, Research and Technology.

The fraud squad in Rennes this month submitted an 'information file' (*'dossier de renseignement'*) to M. Tremoureux, the public prosecutor of the region, making the case for opening an investigation into allegations of fraud at the Laboratory of Nutrition, Lipoprotein Metabolism and Atherosclerosis, at the University of Rennes 1.

The laboratory is headed by Bernard Bihain, and is part of Inserm, the national biomedical research agency. The allegations centre on published claims by the laboratory to have identified and cloned the so-called 'lipolysis stimulated receptor', a molecule involved in fat degradation. The work has been patented by Inserm and the French biotechnology company Genset (see *Nature* 391, 519 & 825; 1998).

The decision on whether to launch a judicial inquiry is the responsibility of the public prosecutor. But police sources point out that 'whistleblowers' who have already testified in the case have made "clear declarations" alleging scientific fraud — a misdemeanour in France — and related offences.

The sources add that knowingly applying for a patent on the basis of fraudulent claims would also be an offence. Olivier Schnerb, the lawyer representing Bihain, last week refused to comment on the grounds that the investigation was continuing.

Further pressure for a judicial inquiry is likely to come from Jacques Lenfant, the president of the University of Rennes 1, and Jean-Pierre Brun, vice-president for research, who intend to invoke a law that requires civil servants to alert the public prosecutor if in the execution of their duties they become aware of a misdemeanour or crime.

In an outspoken article in *Le Monde* last week, Lenfant and Brun also challenged the ministry's handling of the affair. They argued that this fell short of Germany's handling of recent allegations of misconduct, where procedures were established to "immediately treat, drastically and with clarity, suspicions of scientific fraud" (see *Nature* 390, 430; 1997).

The article even alleges that Inserm "has refused a request for scientific clarification... by hiding behind the screen of internal evaluation procedures, which in this case are themselves highly questionable". Ministry

officials could not be reached for comment last week.

The allegations of fraud have been described in particular in a report by a four-member committee of inquiry, chaired by Pierre Corvol, professor of experimental medicine at the Collège de France, in Paris. The committee heard testimony from 24 researchers and technicians who work or have worked in the laboratory.

The Corvol inquiry concluded that the testimonies of seven of these individuals raised doubts about the "validity of certain results published or in the process of publication by the director of the laboratory [Bihain]".

But controversy surrounds the Corvol report. It was not made available to a meeting in November 1997 of the scientific board of Inserm at which Bihain's unit was evaluated.

More extensive investigation called for in the Corvol report was carried out by Bernard Bigot, the then director general of the research ministry. Bigot's report, submitted to the ministry in December, cleared the laboratory of misconduct.

In their article, Lenfant and Brun ask why the minister and his advisers allowed Bigot to "hush up" the Corvol report. Lenfant and Brun describe it as "unacceptable" that the university was not asked to carry out an inquiry, and complain that the university has suffered from "a total absence of information from the ministry".

Daniel Nahon, Bigot's successor at the ministry, announced in May that four international experts had been commissioned to complete a new inquiry within three months (see *Nature* 393, 203; 1998). But this has done little to dampen suspicions. Many remain concerned that the inquiry will only consider published data, and not hear evidence on the conduct of experiments from the 'whistleblowers'.

Some scientists are critical of the role of the media, including *Nature* which first brought the affair into the public eye. Alluding to the role of the press, Marc Pescanski, the Inserm representative on the Corvol committee, deplores "the complete destruction of an Inserm unit despite any evidence of wrongdoing, and before any formal accusation has been made".

Claude Griscelli, the director general of Inserm, last week told *Nature* that the laboratory is to close, as Bihain had asked to be discharged from his duties on "personal grounds". But several scientists nonetheless argue that this should not be used to justify any reduction in efforts to get to the bottom of the affair.

Declan Butler

France sets high targets for impact factors and patents

[PARIS] Claude Allègre, France's minister of national education, research and technology, last week set the target of doubling the impact of French publications in the international scientific literature within four years. He also wants to triple the number of international patents held by French scientists, and to create several hundred stable high-technology companies.

Allègre set these benchmarks following an interministerial meeting on research, chaired by Lionel Jospin, the prime minister. Allègre reiterated the strategy of his ministry, which is to support investigator-driven fundamental research, while exerting control over research with socioeconomic goals.

A National Science Council will be created to advise the government on research strategy. It will be composed of 20 or so eminent scientists and chaired by Allègre. Further advice will be solicited from the Academy of Sciences and other committees (see *Nature* 394, 9; 1998).

Allègre restated that the emphasis of evaluation procedures should shift away from evaluations of laboratories by commissions towards a system based on competitive proposals from individual research groups. More foreign researchers will be involved in conducting evaluations.

To help direct research towards wealth creation, Allègre confirmed the creation of a National Network of Technological Research, which will regroup laboratories working in key economic sectors.

Allègre rejected creating a full-blown postdoctoral system for French fundamental research scientists, and defended France's life tenure system. He said the bigger question was whether scientists should be full-time researchers for the whole of their careers or switch in later years to teaching, for example.

Some scientists say they hoped to see more concrete measures, but a ministry official refutes this criticism, arguing that the government has opened various avenues and that implementing change takes time.

Henri-Edouard Audier, a member of the board of the national union of scientific researchers, gives Allègre the benefit of the doubt, as the ministry now seems more willing to consult interested parties.

But many doubt the government's ability to meet its targets. French research publications have an impact score of 0.9, according to the Paris-based Observatoire des Sciences et des Techniques. Doubling this means overtaking the United States (1.4) and Switzerland (1.5) at the top of the world league. Allègre's claims are "totally unrealistic", says one specialist in scientific indicators. **Declan Butler**

Congress revives hopes for solar power satellites

[WASHINGTON] A congressional committee has recommended a \$20 million increase to the US space agency NASA's budget to study a controversial idea that was all but laid to rest in the 1980s — solar power satellites (SPS) for beaming energy to Earth.

The increase is being urged by members of the House of Representatives, including space subcommittee chairman Dana Rohrabacher (Republican, California), who either have a long-standing interest in solar power satellites or who represent NASA centres that would conduct some of the research.

The idea of collecting solar energy with large satellites in space and transmitting it to the ground via microwaves was first proposed in 1968. Some \$50 million was spent on research at the Department of Energy and NASA in the late 1970s, but the concept was eventually shelved as being too expensive and requiring too many leaps in technology.

Last year, however, a small NASA team completed a two-year study of solar power satellites and found that enough had changed to warrant a revived research programme. Solar cell technology has improved, satellites are smaller and lighter, and SPS concepts no longer need hundreds of astronauts building gigantic structures in space.

NASA had planned to increase spending on SPS-related research from \$2 million to

\$5 million next year. But the Marshall Space Flight Center in Alabama is now working on an implementation plan for SPS research should the House and Senate approve the \$20 million increase later this year.

The House appropriations committee has called for NASA's Marshall and Lewis centres and its Jet Propulsion Laboratory to work with the energy department and private industry on SPS research. NASA is also funding Resources for the Future, a Washington think-tank, to conduct an economic analysis of the concept.

Proponents of SPS are more cautious in selling the idea than in the 1970s, when the satellites "only came in one size, and that was big," according to Whitt Brantley of Marshall's Advanced Systems and Technology Office. Short-term research is likely to focus on modest experiments to advance the basic technology, such as point-to-point transmission of large amounts of microwave energy on the ground or in low orbit.

But SPS must overcome the stigma resulting from the unrealistic plans of 20 years ago. Christopher Flavin of the World-watch Institute, which has been studying the terrestrial solar power industry, calls SPS "a real throwback," and criticizes Congress for starving conventional solar power research while pushing SPS. **Tony Reichhardt**

Accelerator project digs itself into a hole

[MOSCOW] The head of a leading Russian physics institute is under attack from colleagues for refusing to salvage a multi-million-dollar boring machine being used to dig a tunnel for a new particle accelerator.

Anatoly Logunov, director of the Russian Academy of Sciences' Institute of High Energy Physics in Protvino, 70 kilometres south of Moscow, is known for rejecting Einstein's theory of General Relativity, the existence of black holes and the 'Big Bang' for his own theories of matter.

He has decided to abandon a tunnel-boring machine built by the Canadian company Lovat, said to have cost almost US\$18 million, following its completion of a tunnel required for the proton accelerator-accumulator (PAA).

The machine, which is 55 metres long and weighs 360 tonnes, can cover 100 metres in 24 hours, and has almost completed its task. "It would cost only 2.5 million roubles (\$400,000) to bring the machine to the surface," says Vasily Pasika, the engineer-in-chief of the PAA tunnel-building company.

The PAA project received priority

treatment from the government, but has run into financial problems. It was allocated 56 million roubles in the 1998 budget, but in April this was cut to 39 million roubles, and by June the institute had received only 6 million roubles.

Logunov says the institute has no market for the project's only product — mu mesons — and no other income. "Many of our scientists have gone abroad, and with salaries here at less than \$100 a month, I can't blame them," he says. He says the tunnelling machine is the best in the world, "but we have no money to maintain it".

Water enters the tunnel at a daily rate of 30–40 cubic metres, but the pumps are switched off because there is no money for electricity. Technicians, who have long been unpaid, occasionally struggle through the tunnel to check that the borer is working.

There are many potential purchasers for the machine, once it has been brought above ground. It is wanted for the Moscow subway, and four Russian cities — Chelyabinsk, Ufa, Kazan and Krasnoyarsk — have each offered \$18 million for it. **Carl Levitin**

Hungary pushes for further reform

Alison Abbott

Hungarian science has undergone considerable post-communist reform, but the country's academy of sciences feels that further change is needed to allow it to take tough decisions on spending priorities.

[BUDAPEST] The Hungarian Academy of Sciences has asked the government to modify four-year-old legislation liberalizing its operating procedures to introduce more hierarchical forms of decision-making.

At the same time, OTKA, Hungary's main grant-giving body for basic research, plans to introduce more stringent selection procedures, a move that will result in more grant applications being rejected.

Both moves reflect attempts by Hungarian academics to fine-tune hastily introduced post-communist academic reforms because some aspects have proved 'too democratic' to allow for tough but necessary decisions on research priorities and spending.

Although most academics favour a tightening up of the reforms, some fear that reconcentration of power and resources could be bad for Hungary's cash-starved research environment by removing funding from those whose research is not considered an immediate priority.

The 1994 act governing the regulations of the academy, which, like the previous 'Soviet' model, runs a series of basic research institutes and acts as a learned society, was a compromise between two camps.

One opinion, supported by academics jealous of the privileges enjoyed by the academy in communist times, wanted to end its function as a learned society. The other camp, supported mostly by academy staff, opposed such a dramatic change in the 170-year-old institution.

The reformed academy remains a learned society but has broadened its membership to include all holders of PhDs, first introduced in Hungary in 1993 (*Nature* 367, 588; 1994), so university researchers are now included.

A 675-strong decision-making general assembly seems to be working well. But there have been problems with a second body, the Council of the Academic Research Institutes (AKT), which oversees the running of the academy's institutes and the distribution of its HUF5-billion (US\$23 million) budget. Priority-setting has been hampered by conflicts of interest because the AKT includes representatives from all the institutes.

Last year, for example, to reach agreement on tough moves to reform the academy (see *Nature* 386, 426; 1997), the general assembly had to bypass the AKT and create a

Budget rise promise

The painful post-communist contraction of Hungarian science has now stopped. The Hungarian Academy of Sciences was given small budget increases in 1997 and 1998, partly in acknowledgement of reform that has halved its scientific staff since 1990.

Rapid expansion has been promised by the new centre-right coalition government, which took office last month. The largest party, the Young Democrats, promised to double total research spending "in the mid-term", starting by raising it from just under 0.7 per cent to 1.0 per cent of gross national product within the next year.

But many regard these plans, which also include promises to abandon newly introduced student fees, as little more than good intentions with little basis in experience.

Like many in the new government, both the prime minister, Viktor Orbán, and the education minister, Zoltán Pokorni, are in their mid-thirties and have never served in political office before.

A.A.

'consolidation' committee consisting of top academy officials and the heads of its three main scientific committees.

Eventually, around 100 of the academy's scientific posts, and many more administrative and support posts, were abolished, and many institutes were forced to merge.

The parliamentary proposal, drawn up by the academy's presidium and approved by its general assembly after a hard fight, would

replace the AKT with a committee of similar composition to the consolidation committee. It would also increase the power of the academy's elected president.

A newly organized group called the Forum of Scientists, consisting of elected representatives from academy institutes and research groups, last month presented the academy with a counterproposal, retaining the AKT and giving the forum a formal voice in decision-making.

"We decided eight years ago that the academy should have a democratic form of leadership," says György Kéri, a forum member who is director of an academy biology research group based at the Semmelweis University in Budapest. Kéri argues that researchers should take part in decision-making. But the academy's deputy secretary-general, Gábor Náray-Szabó, says the academy is sticking to its original proposal.

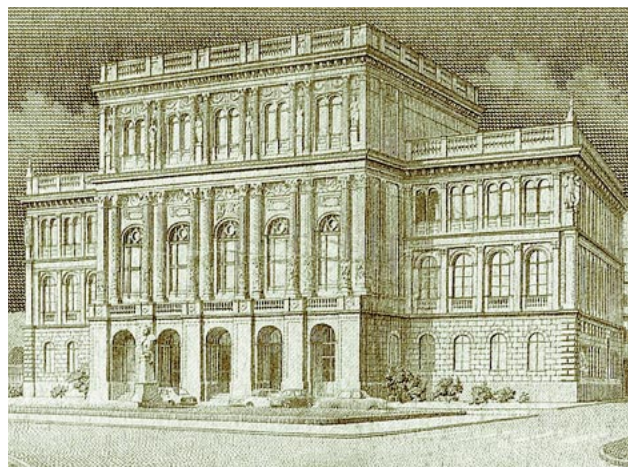
Like the academy, OTKA faces tough decisions about how to distribute its budget. This has halved in value since 1992, and now has to cover both university and academy scientists.

OTKA has set up a solid peer-review process, and has recently established initiatives giving more fundamental support to Hungary's research system. These include a grants programme designed to bring post-docs home from abroad, and a special fund for equipment and journals.

András Lipták, OTKA's president and a professor of biochemistry at the University of Debrecen, says OTKA plans to reduce the number of grants it awards and increase their value; at present, the highest grant is HUF1.5 million and the minimum — which most receive — is HUF500,000.

Lipták says the best researchers must be given enough money to work competitively. But he acknowledges that "some OTKA committees are unhappy about this".

Attila Meskő, a professor of geophysics at the Loránd Eötvös University in Budapest and an OTKA officer, feels that OTKA should distribute small grants more widely to give "moral support to Hungary's many struggling researchers". □



Building for the future: the Hungarian Academy of Sciences wants to change legislation to streamline the decision-making hierarchy that divides up its research budget.

Senate holds up the appointment of US science policy chief

[WASHINGTON] Senate confirmation of the appointment of Neal Lane, the director of the National Science Foundation (NSF), as head of the Office of Science and Technology Policy (OSTP) in the White House could be delayed until next year because of a growing rift between the Senate and the Clinton administration.

Senator Trent Lott (Republican, Mississippi), majority leader in the Senate, said last week that all administration nominations for posts at the White House and at certain other agencies would be denied until the administration cooperated more fully with his investigation into the alleged technology transfer risk of allowing Chinese rockets to launch US satellites. Lott has also said that all nominations will be blocked until after November's Congressional elections unless the administration cooperates with Republican legislation plans.

Rita Colwell, the University of Maryland marine biologist whose nomination to succeed Lane has already been confirmed, does not plan to take over at NSF until Lane moves on.

Japan will help China to clean up power stations

[LONDON] Japan's government is to help reduce industrial pollution in China by providing 1,000 Chinese coal-fired power plants with equipment that will help reduce sulphur emissions.

Officials at Japan's Ministry of International Trade and Industry (MITI) say the decade-long assistance will contribute significantly to pollution control. Three-quarters of China's energy comes from coal. The equipment will be paid for through a combination of loans and official development assistance. MITI will present its plan to China's president Jiang Zemin during his visit to Japan in September.

Euro parliament seeks ban on embryo research

[PARIS] The European Parliament has called for a ban on European Union research involving human embryos. In an amendment to the proposal for the fifth Framework research programme, the parliament calls for existing bans on funding for research on germline gene therapy and human cloning to be extended to include a ban on all research that results in the destruction of human embryos.

Until now, the European Union has avoided this thorny issue, as individual

member states have varying policies reflecting their cultural traditions. The Framework programme is currently the subject of negotiations between the parliament and the European Commission. Meanwhile, the European Group on Ethics in Science and New Technologies, which advises EU institutions on ethical matters, has been asked to give its opinion on the ethics of research on human embryos.

Climate change research wins \$2m in California

[LONDON] The Electric Power Research Institute (EPRI), an organization funded by the US energy industry and based in California, has been given a \$2.2 million award to assess the impact of global climate change in its home state. The award was made by the California state Energy Commission. An EPRI statement said that the organization and its research partners already plan to spend \$30 million on climate change research over the next two years.

UK eyes nuclear power for greenhouse cuts

[LONDON] Britain's Royal Society and Royal Academy of Engineering have launched a study to assess the potential contribution of nuclear energy in reducing greenhouse gas emissions.

The study will be conducted by a 10-member group of experts drawn from industry and the academic community. The group will be chaired by Sir Eric Ash, former rector of Imperial College London, and treasurer of the Royal Society. Its final report will appear next year.

Clinton threatens to veto House funding bill for NIH

[WASHINGTON] President Bill Clinton last week said he will veto a 1999 spending bill passed by the House Appropriations Committee that contains a 9.1 per cent increase for the National Institutes of Health (NIH).

The bill, which funds the Departments of Labor, Health and Human Services and Education, was passed on 14 July on a party-line committee vote. It would boost the NIH budget by \$1.24 billion, to \$14.86 billion. But it contains \$2 billion in cuts to social and education programmes to which the president objects strongly (see *Nature* 394, 108; 1998).

Clinton said in a statement that if he receives the bill in its current form he would have no choice but to veto it.

The Senate has yet to write a corresponding bill. Once it does, the two versions must be reconciled and passed into law by 1 October, the start of the fiscal year.

Manslaughter charges in French blood scandal

[PARIS] France's long-running scandal over contaminated blood took a dramatic turn this week when the commission of investigation of the Court of Justice of the Republic rejected a plea by the French attorney general to drop all charges against three former ministers. The ministers stand charged in connection with their responsibilities in introducing measures to prevent transmission of HIV in blood supplies in the early 1980s.

Instead, the commission sent Laurent Fabius, a former socialist prime minister, and Edmond Hervé and Georgina Dufoix — at the time secretary of state for health, and minister of social affairs, respectively — before the Court of Justice of the Republic on charges of "manslaughter" and "involuntary harm to the integrity of persons".

The charges replace an earlier one of "collusion in poisoning". This was considered by the Supreme Court of Appeals to be inappropriate on the grounds that guilt of this nature required demonstrating a deliberate intention to kill.

Pakistan may sign up to nuclear test ban treaty

[LONDON] Pakistan has taken a small step towards signing the Comprehensive Test Ban Treaty (CTBT) by announcing that any decision would be taken regardless of whether India signed. In the past, Pakistan has made its signature of the treaty conditional on India signing first.

Reports from Islamabad suggest that Pakistan may be prepared to sign the treaty, but will wait for the Indian response before ratifying it. Reports from New Delhi suggest that India may be prepared to sign if it is given nuclear weapons status, and a seat on the United Nations Security Council.

Mathematician drowns attempting record swim



[LONDON] Sir James Lighthill (left), the former Lucasian Professor of Mathematics at the University of Cambridge, died last week attempting to repeat a previous feat: swimming around the

English Channel Island of Sark. He was 74. Lighthill was the first man to perform the nine-hour, nine-mile swim in 1973, at the age of 49.

An authority on fluid dynamics, Lighthill was a former director of the Royal Aircraft Establishment in Farnborough.

Leading scientists still reject God

Sir—The question of religious belief among US scientists has been debated since early in the century. Our latest survey finds that, among the top natural scientists, disbelief is greater than ever — almost total.

Research on this topic began with the eminent US psychologist James H. Leuba and his landmark survey of 1914. He found that 58% of 1,000 randomly selected US scientists expressed disbelief or doubt in the existence of God, and that this figure rose to near 70% among the 400 “greater” scientists within his sample¹. Leuba repeated his survey in somewhat different form 20 years later, and found that these percentages had increased to 67 and 85, respectively².

In 1996, we repeated Leuba’s 1914 survey and reported our results in *Nature*³. We found little change from 1914 for American scientists generally, with 60.7% expressing disbelief or doubt. This year, we closely imitated the second phase of Leuba’s 1914 survey to gauge belief among “greater” scientists, and find the rate of belief lower than ever — a mere 7% of respondents.

Leuba attributed the higher level of disbelief and doubt among “greater” scientists to their “superior knowledge, understanding, and experience”². Similarly, Oxford University scientist Peter Atkins commented on our 1996 survey, “You clearly can be a scientist and have religious beliefs. But I don’t think you can be a real scientist in the deepest sense of the word because they are such alien categories of knowledge.”⁴ Such comments led us to repeat the second phase of Leuba’s study for an up-to-date comparison of the religious beliefs of “greater” and “lesser” scientists.

Our chosen group of “greater” scientists were members of the National Academy of Sciences (NAS). Our survey found near universal rejection of the transcendent by NAS natural scientists. Disbelief in God and immortality among NAS biological scientists was 65.2% and 69.0%, respectively, and among NAS physical

Table 1 Comparison of survey answers among “greater” scientists

Belief in personal God	1914	1933	1998
Personal belief	27.7	15	7.0
Personal disbelief	52.7	68	72.2
Doubt or agnosticism	20.9	17	20.8
Belief in human immortality	1914	1933	1998
Personal belief	35.2	18	7.9
Personal disbelief	25.4	53	76.7
Doubt or agnosticism	43.7	29	23.3

Figures are percentages.

scientists it was 79.0% and 76.3%. Most of the rest were agnostics on both issues, with few believers. We found the highest percentage of belief among NAS mathematicians (14.3% in God, 15.0% in immortality). Biological scientists had the lowest rate of belief (5.5% in God, 7.1% in immortality), with physicists and astronomers slightly higher (7.5% in God, 7.5% in immortality). Overall comparison figures for the 1914, 1933 and 1998 surveys appear in Table 1.

Repeating Leuba’s methods presented challenges. For his general surveys, he randomly polled scientists listed in the standard reference work, *American Men of Science* (AMS). We used the current edition. In Leuba’s day, AMS editors designated the “great scientists” among their entries, and Leuba used these to identify his “greater” scientists^{1,2}. The AMS no longer makes these designations, so we chose as our “greater” scientists members of the NAS, a status that once assured designation as “great scientists” in the early AMS. Our method surely generated a more elite sample than Leuba’s method, which (if the quoted comments by Leuba and Atkins are correct) may explain the extremely low level of belief among our respondents.

For the 1914 survey, Leuba mailed his brief questionnaire to a random sample of 400 AMS “great scientists”. It asked about the respondent’s belief in “a God in

intellectual and affective communication with humankind” and in “personal immortality”. Respondents had the options of affirming belief, disbelief or agnosticism on each question¹. Our survey contained precisely the same questions and also asked for anonymous responses.

Leuba sent the 1914 survey to 400 “biological and physical scientists”, with the latter group including mathematicians as well as physicists and astronomers¹. Because of the relatively small size of NAS membership, we sent our survey to all 517 NAS members in those core disciplines. Leuba obtained a return rate of about 70% in 1914 and more than 75% in 1933 whereas our returns stood at about 60% for the 1996 survey and slightly over 50% from NAS members^{1,2}.

As we compiled our findings, the NAS issued a booklet encouraging the teaching of evolution in public schools, an ongoing source of friction between the scientific community and some conservative Christians in the United States. The booklet assures readers, “Whether God exists or not is a question about which science is neutral”⁵. NAS president Bruce Alberts said: “There are many very outstanding members of this academy who are very religious people, people who believe in evolution, many of them biologists.” Our survey suggests otherwise.

Edward J. Larson

Department of History, University of Georgia,
Athens, Georgia 30602-6012, USA
e-mail: edlarson@uga.edu

Larry Witham

3816 Lansdale Court, Burtonsville,
Maryland 20866, USA

1. Leuba, J. H. *The Belief in God and Immortality: A Psychological, Anthropological and Statistical Study* (Sherman, French & Co., Boston, 1916).

2. Leuba, J. H. *Harper's Magazine* **169**, 291–300 (1934).

3. Larson, E. J. & Witham, L. *Nature* **386**, 435–436 (1997).

4. Highfield, R. *The Daily Telegraph* 3 April, p. 4 (1997).

5. National Academy of Sciences *Teaching About Evolution and the Nature of Science* (Nat'l Acad. Press, Washington DC, 1998).

Dilemma over genetics and population in China

Sir—As you write in a recent editorial, the forthcoming 18th International Congress on Genetics in Beijing will provide a rare opportunity to continue discourse on the ethics and science of eugenics¹. But a number of fundamental problems facing geneticists in China deserve mention. Without a thorough understanding and awareness of these problems and their cultural and psychological roots — as

eloquently expounded recently by Chen-Lu Tsou² — the discourse may not produce significant tangible results.

We acknowledge that the Chinese Maternal and Infant Health Law, which has been the focus of much recent debate, represents a well-intentioned step towards reducing the burden of debilitating diseases perceived to be hereditary. By sheer scale, the enormous social and economic cost to the most populous nation has no equal in the world, and would surely prompt any sensible society to react. Indeed, given the urgency of the population problem

confronting the nation, it may appear logical to concentrate on the segments of its society considered the least productive and least able to contribute to the future, if a major effort is launched to reduce its population size³.

But the good intention of the law is seriously undermined by its shaky scientific foundations. For example, where is the evidence that 20 million people are handicapped by hereditary diseases? In a nation where more than half the adult male population smoke, and environmental pollution is rampant in some areas, could a

significant proportion of those presumed hereditary handicaps be prevented by a reduction in smoking, a cleaner environment, and improved pre-, peri- and postnatal care? Should one take the current search for genetic mechanisms underlying many complex diseases or disorders, such as schizophrenia, as the *fait accompli* that these diseases are preventable through sterilization? How strong is the evidence that enforcement of the law alone will prevent many or all of the handicaps? Without solid documentation, any claims about the law and its intended effects are merely opinions, without scientific validation.

The fact, as pointed out by the sponsor of the law, that births of "inferior quality" are relatively more common among "the old revolutionary base, ethnic minorities, the frontier, and economically poor areas"⁴ suggests that many so-called "inferior births" may in fact be of environmental origin, and so preventable through improved living standards and better pre-, peri- and postnatal care (for example, taking folic acid, reducing perinatal trauma, and eliminating iodine deficiency).

The law was drafted with input from geneticists in China, but it is questionable whether the scientific part of the law was based on the best knowledge available. Judging from Chinese human genetics textbooks and scant publications in international journals, it is evident that basic research in genetic epidemiology is still in its infancy in China. This situation is undeniably the result of political turmoil and the chronic shortage of government funds for this type of research.

The lack of a rigorous grant review system allows scope for excessive importance to be given to popular acclaim and to the political goals of scientific research in allocating funding. In a country where political loyalty is often considered more important than scientific talent and integrity, this can be an effective strategy for attracting government funds.

Indeed, the recent substantial increase in funding for genetic research in China⁵ was largely the result of a letter to the Chinese president from a prominent geneticist urging protection of China's human genetic resources, because of the fear of losing the resources to foreign organizations.

As for issues relating to international collaboration, the importance given to popular acclaim and political goals can easily lead to narrow-minded nationalism, which can be generated by provoking painful memories of imperialist aggression and humiliation in the past. This type of

nationalism, coupled with the lack of 'checks and balances' in the system, provides a recipe for abuse. Ironically, efforts to protect China's human genetic resources are seriously compromised by inadequate research in basic genetic epidemiology.

The best protection against over-politicization and ignorance may be an overhaul of China's research evaluation and grant review systems. For example, experts from other countries should be invited to participate in evaluating large scientific grant applications and research institutions. For a poor country such as China, this is also the best way to ensure that scarce and meagre resources are well spent. The science part of the eugenics law — which unfortunately has no quick fix — requires years of basic genetic research which will ultimately benefit not only the Chinese people, but all humankind.

The opinions expressed here are the authors' own and should not be taken to represent those of their institutions.

Sun-Wei Guo

*Institute of Human Genetics and
Division of Epidemiology,
University of Minnesota,
1300 South Second Street, Suite 300,
Minneapolis, Minnesota 55454-1015, USA
e-mail: swguo@med.umn.edu*

Chang-Jiang Zheng

*Epidemiology, Statistics, Data System Branch,
National Institute on Deafness and Other
Communication Disorders,
Bethesda, Maryland 20892, USA*

C. C. Li

*Department of Human Genetics,
University of Pittsburgh, Pittsburgh,
Pennsylvania 15261-0001, USA*

1. *Nature* **392**, 109 (1998).
2. Tsou, C.-L. *Science* **280**, 528–529 (1998).
3. Neel, J. V. *Perspect. Biol. Med.* **40**, 328 (1997).
4. Dickson, D. *Nature* **367**, 3 (1994).
5. Swinbanks, D. *Nature* **394**, 109 (1998).

Science in the firing line in Argentina

Sir — I am appalled by your editorial of 16 April, asserting that Argentina's national research council (CONICET) spends nearly all its money on 3,000 staff scientists but has now set up a new agency, in collaboration with industry, to produce sound science (*Nature* **392**, 635; 1998).

I would like to mention a few facts and events that influenced the careers of those 3,000 researchers. In 1961, a secretary of state who did not tolerate Jews fired the director of the Malbran Institute of Immunology, Ignacio Pirofsky, provoking

the resignation and exile of most of the research personnel. In 1966, the military destroyed the School of Exact and Natural Sciences of the University of Buenos Aires ("Exactas"), beating and incarcerating researchers, and had the laboratories exorcised by a priest. Luis Botet was appointed president of the university by the military: he fired all scientists who expressed solidarity with their Exactas colleagues. Hence, 1,315 scientists left the country.

In 1976, Raul Matera, undersecretary of science and technology, bought 40 crucifixes for CONICET's offices, despite the small amount of money available for research. The Argentinian government passed a law of amnesty to pardon all the military involved in torturing and murdering tens of thousands of people, as well as a law of *punto final* (no more questions asked), but refused to compensate researchers who had been fired and deprived of their labs. In 1990, finance minister Domingo Cavallo said that he would prefer scientists "to wash dishes". Accordingly, those 3,000 researchers referred to in your editorial are paid meagre salaries, and have almost no money to run their labs. Argentina has a far larger and more productive community of researchers in exile abroad than it has at home.

To appreciate the quality of Argentinian researchers, one has only to note that they publish in the best international journals, they frequently work in first-rate universities in Britain, France and the United States, and are awarded all types of distinctions, including the Nobel prize. Argentina has some poorly financed research, but no science, because while the first depends on the ability of a few thousand, science is a way of interpreting reality that Argentina has never developed. Thus, not a single workers' union or society of entrepreneurs complained when the universities were destroyed. Consequently, today masses of unemployed people beg San Cayetano (the patron saint of workers) for work.

Argentina is not willing to accept that, in order to develop science, an ethical transformation is required, not just of its scientific infrastructure but of its society in general. Its ideal seems to be to combine technology with theology. For local governments, science is something that comes only after countries become rich. So they appoint managers to decide scientific matters.

Marcelino Cerejido

*Centro de Investigación y de Estudios Avanzados,
Fisiología, Biofísica y Neurociencias,
Apartado Postal 14-740,
07000 Mexico
e-mail: cerejido@laneta.apc.org*

Dolly is a clone — and no longer alone

Davor Solter

After the furore that surrounded the arrival of Dolly, the first mammal to be cloned from differentiated adult cells, doubts were raised that she really was a clone. Those doubts can now be set aside, and the technique has been further validated by the cloning of mice.

Two years ago, a report¹ that lambs had been successfully cloned from cultured, differentiated cells (that is, cells committed to be just one particular type) passed relatively unnoticed. Just one year later, the world sat up and took notice when the cloning of cells derived from an adult udder resulted in the birth of Dolly². Subsequent discussions revealed the depth (or shallowness) of our homocentricity — very little was said about what we can learn from cloning experiments. This misplaced emphasis led to repeated statements, culminating in a letter by Sgaramella and Zinder³, that a single case of cloning from an adult cell would have to be repeated before it could be taken seriously. Moreover, they suggested that Dolly might actually have been cloned from a fetal cell that had contaminated the udder-cell culture through a biological or laboratory accident.

Reports elsewhere in this issue now lay these doubts to rest. By throwing the full panoply of forensic DNA-testing methods at the problem of Dolly's origin (including the sworn, unbroken chain of custody in providing samples), Ashworth *et al.*⁴ and Signer *et al.*⁵ (page 329) have shown that Dolly is indeed the direct descendant of an udder cell derived from a nameless Finn Dorset ewe. They compared DNA isolated from frozen udder tissue, the cell culture derived from it and Dolly's blood, and found that the three DNA samples were identical. Further comparisons with DNA from control sheep indicate that the probability of contamination with fetal cells or mistakenly mixed cell cultures is vanishingly small.

But the fact of Dolly's uniqueness remains. Thus the significance of the report by Wakayama *et al.*⁶ (page 369), who describe the existence (if my maths is correct) of 22 healthy, cloned female mice. In each case, the nuclear material was removed from a mouse egg cell and replaced by the nucleus from a granulosa cell (a differentiated type of cell that surrounds the egg), in a process known as nuclear transfer (see Fig. 5 on page 372). The importance of this report cannot be overemphasized — mice have a short gestation period, well-characterized genetics, and their embryos are much easier to manipulate than those from larger mammals, opening

up the possibility of a broad experimental analysis of mammalian cloning and of the factors that determine its outcome.

Probably spurred on by the successes in cloning large farm animals, Wakayama *et al.* ignored the series of reports (my own included⁷) that stressed the difficulties — if not impossibilities — of cloning mice by simple nuclear transfer. One barrier to cloning was thought to be the early activation, during development, of the mouse embryonic genome, thus leaving too little time for the transferred nucleus to be reprogrammed. Normally, every type of differentiated cell expresses its own unique set of genes. When, during cloning, the nuclear material is transferred to the egg from a differentiated cell, it needs to be reprogrammed so that it can give rise to not just one type of cell, but to all of the different cells that make up a mouse. Reprogramming is thought to involve the shutting down of all (or most) of the genes, before the new sets are switched on. Wakayama and colleagues have now shown that there is, in fact, enough time for reprogramming to occur, and they deserve every praise for their skill and perseverance.

Given that so many of us failed, it is not immediately clear why Wakayama *et al.* have succeeded, although they list several technical reasons that need to be explored further. The point is, now that the cloning of mammals has become an accomplished fact and an accessible technique (and once the talking heads and presidential commissions have issued their reports and opinions), we can try to set some near and distant goals.

The purpose of a specific cloning experiment will largely be determined by the choice of cloned subject (Fig. 1), but Wakayama and colleagues' results indicate that some adult cells can be cloned whereas others cannot. Is this a biological or technical problem? If biological, what makes some cells reprogrammable and others not? And what is the nature of the reprogramming? Moreover, the success rate of cloning is still relatively low. Does this reflect the heterogeneous nature of the cells that are used to provide the donor nuclei, or is it due to the unpredictable nature of the reprogramming process? We have just started to characterize the genes that are expressed in a stage-specific manner during early mouse development^{8,9}, and these results could be used to determine whether the same set of genes is switched off and on in all of the cloned embryos, or whether each embryo is unique. This may allow us to work out which genes must be turned on and off for nuclear transfer to succeed.

The scope of future basic biological research has been increased by the success of cloning. For example, mammals possess a set of so-called imprinted genes that are expressed depending on which parent they are inherited from. Both paternally and maternally imprinted genes are necessary for normal development. Now, using nuclear

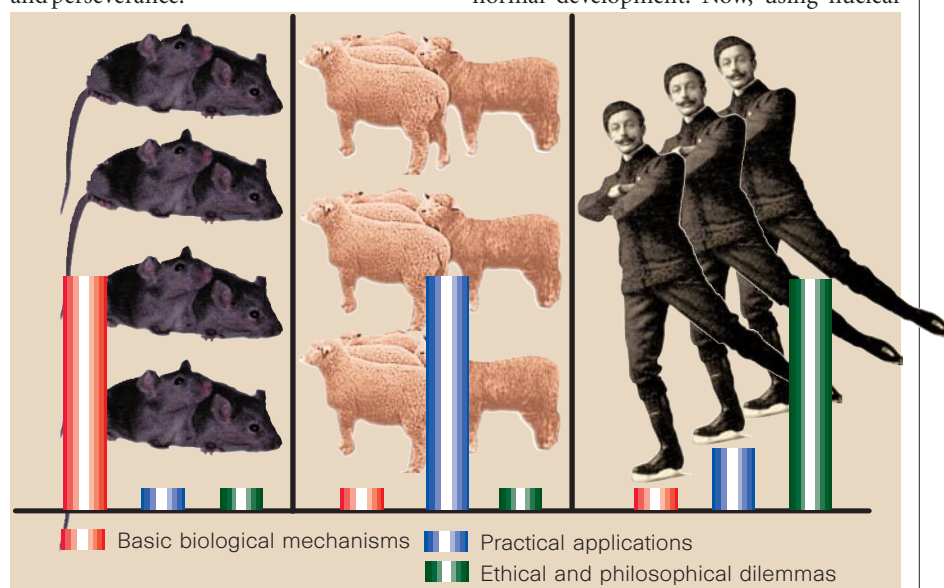


Figure 1 Cloning strategies — the subject to be cloned determines the relative importance of various problems and questions related to cloning. Analysis of basic biological mechanisms will predominate in the cloning of small laboratory mammals; moral and philosophical questions will affect largely the cloning of humans; and practical uses and benefits will determine the extent of cloning of large farm animals and, to a degree, of humans as well.

transfer, we can determine whether imprinting is preserved or erased in differentiated cells. The low success rate of cloning may turn out to be due to a random loss of correct imprinting. Moreover, differentiated female cells contain one inactive X chromosome, whereas in early female embryos both X chromosomes are active. So, we can learn much about inactivation of the X chromosome by monitoring what happens to the inactive X chromosome after nuclear transfer. Cloning by nuclear transfer from established, well-characterized cell lines will be much more informative than cloning from cells that have been freshly isolated from adult or fetal tissues.

With the cloning of large farm animals, the goals become economic¹⁰. The profit motive has, fortunately, kept cloning research alive, despite initial difficulties. Genetically altered fibroblasts (connective-tissue cells) can now be used to clone sheep¹¹ and cows¹² by nuclear transfer, and this should allow us to engineer the large-scale production of useful proteins by farm animals. Obviously, cultured cells must be used for precise targeting of the desired genetic manipulation, be it the addition or deletion of a specific gene. This again highlights the need to establish well-defined, cultured cell lines to be used for cloning.

Finally, what about cloning humans? At some point we will have to determine whether and when cloning — in the sense of taking somebody's cell nucleus, transferring it into an egg and raising the embryo to term and beyond — should be attempted. But we must remember that cloning is not an instant duplication, so mad dictators will not be able to expand themselves into huge armies of doppelgängers, nor will the bereaved be able to restore their lost ones. There is, nevertheless, one area in which cloning technology can be useful to humans: cell and tissue therapy.

Practical problems notwithstanding, at present there are no theoretical obstacles to such tissue therapy. Embryonic stem cells have the ability to differentiate into any cell type and could be produced from human blastocysts (embryos at a very early stage of development). Indeed, this has been done repeatedly in mice. This means that people could provide their own cells and, by using them to replace the nuclei of their own or donor eggs, obtain stem cells in culture. These cells could then be induced to differentiate in culture, providing individually tailored cell and tissue replacements without the current problems of rejection that affect transplantation from the same or foreign species. But, as far as I know, it would be illegal in most countries to sacrifice such blastocysts by turning them into a cell line in culture, because they represent potential human beings. Moreover, many technical problems will have to be over-

come before this approach can become reality, but its potential value for human medicine is enormous.

There are many other beneficial aspects of cloning — easy preservation of genetically important strains and mutants of laboratory and farm animals, preservation and propagation of endangered species (provided that interspecies nuclear transfer is biologically possible) and genetic improvements, among others. There are undoubtedly many obvious and hidden dangers as well, although these do not include the groundless fears and blanket bans that seem to be the knee-jerk reaction to any cloning news. The work of Wakayama and colleagues⁶ may be the last report (or at least the penultimate, as the first cloning of humans is bound to raise a few eyebrows) to induce a media frenzy. After this, I hope that we can move towards times

when cloning will become a standard and useful technique to address numerous problems in basic and applied science. □

Davor Solter is in the Department of Developmental Biology, Max-Planck-Institut für Immunbiologie, Stübeweg 51, 79108 Freiburg, Germany.
e-mail: solter@immunbio.mpg.de

1. Campbell, K. H. S., McWhir, J., Ritchie, W. A. & Wilmut, I. *Nature* **380**, 64–66 (1996).
2. Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. S. *Nature* **385**, 810–813 (1997).
3. Sgaramella, V. & Zinder, N. D. *Science* **279**, 635–636 (1998).
4. Ashworth, D. et al. *Nature* **394**, 329 (1998).
5. Signer, E. N. et al. *Nature* **394**, 329–330 (1998).
6. Wakayama, T., Perry, A. C. F., Zuccotti, M., Johnson, K. R. & Yanagimachi, R. *Nature* **394**, 369–374 (1998).
7. McGrath, J. & Solter, D. *Science* **226**, 1317–1319 (1984).
8. Rothstein, J. L. et al. *Genes Dev.* **6**, 1190–1201 (1992).
9. Oh, B., Hwang, S.-Y., Solter, D. & Knowles, B. B. *Development* **124**, 493–503 (1997).
10. Anderson, G. B. & Seidel, G. E. *Science* **280**, 1400–1401 (1998).
11. Schnieke, A. E. et al. *Science* **278**, 2130–2133 (1997).
12. Cibelli, J. B. et al. *Science* **280**, 1256–1258 (1998).

Nonlinear dynamics

Death by delay

Steven H. Strogatz

Anyone who has ever taken a shower in an old dormitory knows that time delays can cause oscillatory instabilities. First the water is too cold, so you turn up the heat. But as you stand there shivering, the water does not get any hotter, because of delays in the antiquated heating system. If you impatiently turn up the heat even more, you are in trouble — by now, your original request for heat has registered and the water is scalding. Furiously reversing the setting, you are about to trigger a series of wild oscillations.

This much is familiar. The new twist, reported in *Physical Review Letters* by Reddy, Sen and Johnston¹, is that delays can also have the opposite effect: they can damp out oscillations that would otherwise be self-sustaining. More precisely, coupled limit-cycle oscillators can drive one another to a state of zero amplitude — often called amplitude death — if their mutual interactions are suitably delayed. This 'death by delay' may affect coupled limit-cycle oscillators in physics, medicine, biology and chemistry.

In the new models, each oscillator is assumed to have a stable limit cycle (Fig. 1). This assumption is appropriate for any system that will oscillate on its own, that is, in the absence of external periodic forcing. Examples include the heart's own pacemaker cells, rhythmically chirping crickets, flashing fireflies, a pendulum clock with an escapement, aeroelastic flutter in airplane wings, oscillating chemical reactions, and Josephson junctions driven by a constant-bias current².

In the earliest research on limit-cycle oscillators, between about 1920 and 1970, most theorists concentrated on the behaviour of a single, periodically forced oscillator, or two coupled oscillators. The questions were prompted in part by such emerging nonlinear technologies as vacuum tubes, phase-locked loops, radar and lasers, as well as by their intrinsic mathematical interest. Nowadays, the important theoretical problems concern the collective behaviour of large arrays of oscillators³. Synchronization, wave propagation, spatial patterns and self-organized criticality have

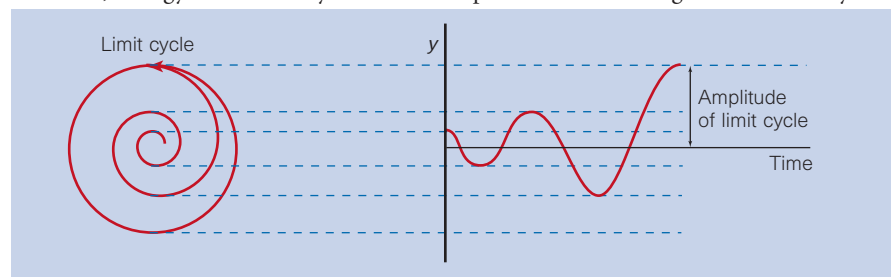


Figure 1 A self-sustained oscillation in voltage, concentration or some other physical variable is represented geometrically by a limit cycle. Here the y-axis records a measurable variable, and the other coordinate represents the remaining variable (such as current, or rate of change of concentration) needed to characterize the state of the system completely. If a disturbance suddenly reduces the amplitude, the oscillation spontaneously builds back up to its standard size.

attracted the most attention, with applications ranging from neurobiology to superconducting electronics.

But the effects of time delay have rarely been considered⁴⁻⁷. That might seem odd. After all, everyone knows that delays are inevitable in any real system, owing to the finite speed of propagation, synaptic processing delays, reaction times and so on. But from a mathematical perspective, coupled nonlinear differential equations are already a headache to analyse; if delay is included as well, the difficulties become downright intimidating.

So Reddy, Sen and Johnston¹ have sensibly chosen to simplify other aspects of their problem. They first consider a pair of limit-cycle oscillators. In the absence of coupling, the state of the system can be visualized as a pair of points P and Q moving on a plane (Fig. 2). Both points are attracted to the same limit cycle, on which they rotate with constant angular frequencies. If perturbed away from the cycle, each point will spiral back to it.

Now suppose the oscillators are coupled. The form of the delayed interaction assumed by Reddy, Sen and Johnston has a nice geometric interpretation. It can be pictured as a force that pulls one point towards the other. But because of the assumed delay in the coupling, each point is not pulled towards the current position of the other, but rather towards the position at a time τ in the past. Figure 2 suggests that for suitable values of τ , the coupling tends to pull both oscillators off the cycle, reducing their amplitudes. In the extreme case, both oscillators could drag each other down to zero amplitude and stay there.

That idea is confirmed by the analytical arguments of Reddy, Sen and Johnston. They show explicitly that amplitude death occurs within certain islands in the parameter space of coupling strength K and delay τ , and they get similar results for arrays of N identical oscillators with delayed mean-

field coupling. (Fortunately, the calculations of island boundaries reduce to linearization about a fixed point, and the resulting linear delay-differential equations are analytically tractable.) Numerical simulations suggest that delay can also cause amplitude death in models with nearest-neighbour coupling.

The main question now is whether death by delay can be observed in experiments. The authors suggest several candidate systems. In my view, the most promising are oscillating chemical reactions in coupled stirred tank reactors^{8,9} and high-power microwave oscillators coupled by signals propagating through the connecting waveguide bridges¹⁰. In both cases, the control parameters can be adjusted so that the oscillators are strongly coupled and weakly nonlinear, and the delays are long enough — three requirements for the theory to apply.

On the other hand, the speculation¹ that death by delay might also be involved in arrhythmias of the cardiac pacemaker network seems far-fetched — the propaga-

tion delays are probably too short and the limit cycles are too strongly attracting for this mechanism to apply. Otherwise, 'death by delay' could take on a grim double meaning. □

Steven H. Strogatz is in the Department of Theoretical and Applied Mechanics, and the Center for Applied Mathematics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA.

e-mail: strogatz@cornell.edu

1. Reddy, D. V. R., Sen, A. & Johnston, G. L. *Phys. Rev. Lett.* **80**, 5109–5112 (1998).
2. Strogatz, S. H. *Nonlinear Dynamics and Chaos* (Addison-Wesley, Reading, MA, 1994).
3. Pérez, C. J., Corral, A., Díaz-Guilera, A., Christensen, K. & Arenas, A. *Int. J. Mod. Phys. B*, **10**, 1111–1151 (1996).
4. Niebur, E., Schuster, H. G. & Kammen, D. *Phys. Rev. Lett.* **67**, 2753 (1991).
5. van Vreeswijk, C., Abbott, L. F. & Ermentrout, G. B. *J. Comp. Neurosci.* **1**, 313–321 (1994).
6. Bressloff, P. C. & Coombes, S. *Phys. Rev. Lett.* **78**, 4665–4668 (1997).
7. Ernst, U., Pawelzik, K. & Geisel, T. *Phys. Rev. E* **57**, 2150–2162 (1998).
8. Bar-Eli, K. *Physica D* **14**, 242–252 (1985).
9. Crowley, M. & Epstein, I. R. *J. Phys. Chem.* **93**, 2496–2502 (1989).
10. Benford, J., Sze, H., Woo, W., Smith, R. R. & Harteneck, B. *Phys. Rev. Lett.* **62**, 969–972 (1989).

Ras signalling

Caught in the act of the switch-on

Fred Wittinghofer

In everyday life we are surrounded by switches that we turn on and off: from the alarm clock (which rings too early) in the morning to the light in the evening. Flipping the switch seems to be such a simple process that we don't think about the underlying mechanism. The cell has molecular switches in the form of proteins, and people have been thinking hard about how they might work. One such switch called Ras has been the subject of particular attention, and from the work of Boriack-Sjodin *et al.* (page 337 of this issue¹) we get a good idea of how it is turned on with the help of another protein, Sos.

Molecular switches such as Ras are GTP-binding proteins², and they cycle between a GTP-binding active on state and a GDP-binding inactive off state. They become active when protein-bound GDP is released and GTP binds instead. The bound nucleotide — triphosphate or diphosphate — determines whether these proteins relay a signal, because downstream molecules only interact with the GTP-bound form. They switch off by splitting bound GTP to GDP (which stays bound) and inorganic phosphate (which is released), and the rate of this reaction determines the duration of the signal. The problem for the cell is that GDP is bound very tightly and dissociates very slowly, yet it needs to be released within seconds so that the protein can be loaded with GTP very quickly. For this, GTP-binding proteins need the help of another class of protein

called guanine-nucleotide-exchange factors (GEFs).

Ras is an especially important GTP-binding protein, because it sits in the plasma membrane and is part of the mechanism for transmitting signals for growth and many other processes from the outside to the inside of the cell. A lot of growth factors activate Ras, which, in a series of chemical reactions and protein interactions, then ends up telling the nucleus to initiate the cell cycle and duplicate its DNA. Failure to deactivate Ras has disastrous consequences: mutants of Ras that cannot be deactivated are an important element in tumour formation and are found in about 30% of human tumours.

Molecular aspects of the Ras switch-off mechanism were unveiled last year³, and now we have a close, three-dimensional look at the switch-on mechanism. A GEF for Ras was first identified in yeast as Cdc25; later, mammalian homologues with a Cdc25-like GEF domain were found, such as Cdc25^{Mm} and Sos (for Son of sevenless, named after a *Drosophila* gene)⁴. Sos is located in the cytoplasm of resting cells. When a cell is stimulated by growth factors, Sos is recruited to the plasma membrane, ready to act on Ras.

The ensuing mechanistic problem is as follows: the nucleotide comes off with a dissociation rate constant of 10^{-5} s, which means that it would take hours. By reducing the affinity appropriately, Sos (or Cdc25) increases the dissociation rate by several orders of magnitude so that the nucleotide

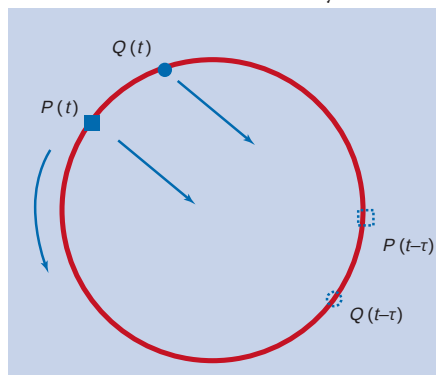


Figure 2 Two limit-cycle oscillators, with a time-delayed coupling¹. The current state $P(t)$ of one oscillator is pulled towards the retarded state $Q(t-\tau)$ of the other, and vice versa. For appropriate values of delay and coupling strength, both oscillations will spiral inwards and die out.

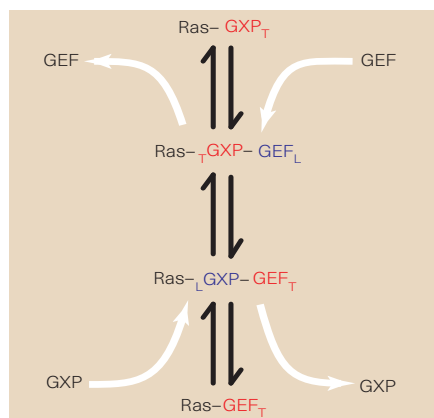


Figure 1 Kinetic scheme for the nucleotide exchange reaction on Ras catalysed by guanine-nucleotide-exchange factors (GEFs); nucleotide is designated as GXP. The Ras–nucleotide binary complex interacts with GEF to form a ternary complex where the nucleotide is still bound tightly (T) and the GEF is bound loosely (L). After a (possibly rate-limiting) conformational change, the ternary complex with loosely bound nucleotide relaxes into a binary Ras–GEF complex. Such a complex — whose structure has been solved by Boriack-Sjodin *et al.*¹ with Son of sevenless (Sos) as the exchange factor — is stable only in the absence of nucleotide. In its presence the reaction is rapidly reversed towards a Ras–nucleotide complex, which is mostly Ras–GTP because of the prevalence of GTP in the cell. In principle, however, GEF also catalyses the conversion of Ras–GTP to Ras–GDP.

comes off in less than a second (as shown for Cdc25; ref. 5). The usual view of events is that Sos approaches the Ras–GDP binary complex and forms a ternary complex (Fig. 1). Thereafter the nucleotide-binding pocket goes from a tight to a loose binding conformation, resulting in a fast release of nucleotide from the ternary Ras–GDP–Sos complex. By binding GTP, the dominant nucleotide in the cell, a new ternary complex is formed and then Sos and Ras–GTP are released (Fig. 1).

Using X-ray crystallography, Boriack-Sjodin *et al.*¹ have solved the three-dimensional structure of the binary complex between Ras and a fragment of Sos. It shows beautifully the tricks Sos uses to kick out GDP. The GEF fragment of Sos is a mostly helical protein consisting of the actual catalytic (Cdc25 homology) domain and the REM (for Ras exchanger motif⁶) domain. This REM domain has been assumed to be required for interaction with Ras, and from the structure it seems that it orientates and/or stabilizes two conspicuously protruding helices (a helical hairpin) of the catalytic domain (Fig. 2). The three-dimensional structure does not resemble those known for other GEFs^{7–10}, implying that the structural mechanism of nucleotide release is different for other switch proteins.

The structure of Ras has been determined

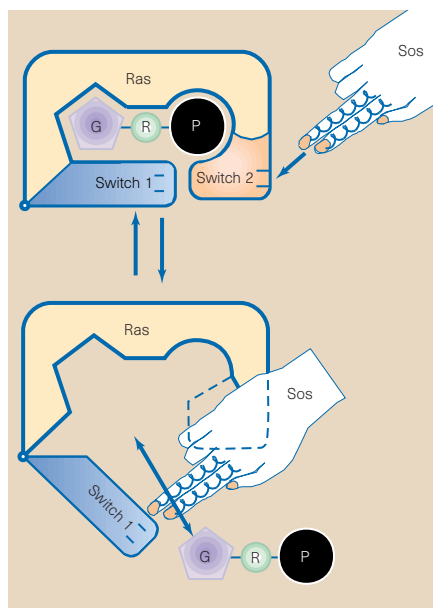


Figure 2 Schematic view of the binary Ras–Sos complex and its implications for catalysis. Nucleotide is bound mostly by the guanine base (G) and phosphate plus Mg^{2+} moiety (P) of the nucleotide, while the ribose (R) does not contribute. In the binary Ras–Sos complex, the binding site of the phosphate moiety is occupied by the interaction of Sos with Switch 2, while the base-binding area is opened up because the Switch 2 region of Ras has been flipped away by the helical hairpin of Sos. (Drawing by G. Horn.)

before, and it showed that Ras has two patches on the surface, called switch regions, that have different conformations in the inactive and the active state. These regions are involved in the molecular interactions of Ras with other proteins, and biochemical studies had shown that GEFs also interact with them. Sos indeed binds directly to the Switch 2 area, blocking the entry of the phosphate moiety of the nucleotide, including the Mg^{2+} ion (which is needed for tight binding), into the binding site (Fig. 2). At the same time Sos uses the helical hairpin to flip away the Switch 1 region in such a dramatic fashion that the base-binding part of the nucleotide-binding site is completely open. In doing so, Sos probably takes advantage of the fact that the switch regions of Ras are very mobile, and that Switch 1 is found in at least two conformations in solution¹¹.

The structure thus tells us why the affinity of the nucleotide is so dramatically reduced. Although it provides just one snapshot in a series of reactions leading to nucleotide exchange (Fig. 1), it suggests how the nucleotide can re-enter its binding site and reverse the whole process. The nucleotide probably comes in with its base moiety first; but to achieve full-strength binding, Ras needs to undergo a conformational change that breaks the interactions with Sos in the phosphate-binding area, releasing Sos and



100 YEARS AGO

Smithson died in Genoa in 1829, having bequeathed all his property to a nephew, Henry James Hungerford by name, and after him to any child of this nephew, “legitimate or illegitimate”; but in case of the said nephew dying and leaving no child, then all the property was, as mentioned above, to go “to the United States of America, to found at Washington, under the name of the Smithsonian Institution, an establishment for the increase and diffusion of knowledge among men.” Henry Hungerford died unmarried and without heirs in 1835, and Smithson’s solicitors forthwith communicated with the United States Embassy in London. Then followed discussions in Senate and House of Representatives. Some senators considered that it would be beneath the dignity of the nation to receive benefits from a foreigner. Other senators considered that it would not. The House of Representatives referred the matter to a select committee, and finally the legacy was accepted, and Richard Rush, a lawyer of high standing, at one time United States Minister at the Court of St. James’s, was selected to prosecute the claim in Chancery. When Mr Rush arrived in London he found that there were eight hundred cases in Chancery ahead of his, yet he managed to get the suit settled in less than two years, a matter “which gave rise to no little surprise,” seeing that “the English lawyers themselves admitted that a Chancery suit was a thing which might begin with a man’s life, and its termination be his epitaph.”

From *Nature* 21 July 1898.

50 YEARS AGO

Mr. J. T. Sorenson included a petrosal [bone] from a blackfish (*Globicephala meloena*) in the collection which he made in his second stay at Campbell Island in the New Zealand subantarctic group. The extraordinary density of this specimen is striking and immediately apparent when the bone is handled. Although the specimen measures only 40 mm. × 28 mm. × 18 mm. in its major dimensions, the dry weight is 17.5 gm. After boiling in water to expel air, the volume measured by immersion is 5.0 c.c.; accordingly, the physical density is 3.5, considerably exceeding the range given in the Smithsonian “Physical Tables”.

From *Nature* 24 July 1948.

allowing the nucleotide to bind tightly. By the principle of microscopic reversibility, we can also deduce the structural changes in the forward reaction.

Much remains to be done to understand each of the partial reaction steps leading to Ras activation, and to relate the kinetics and dynamics to the structural changes of the full reaction. But a giant leap forward has been made and experiments to tackle further problems can now be designed in a rational way. So the next time you use the switch to turn the light on or turn the alarm clock off (because it's too early), remember how clever the cell is in switching on GTP-binding proteins. □

Fred Wittinghofer is in the Max-Planck-Institut für Molekulare Physiologie, Rheinlanddamm 201, 44139 Dortmund, Germany.

e-mail: alfred.wittinghofer@mpi-dortmund.mpg.de

1. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. & Kuriyan, J. *Nature* **394**, 337–343 (1998).
2. Bourne, H., Sanders, D. & McCormick, F. *Nature* **348**, 125–132 (1990).
3. Scheffzek, K. *et al.* *Science* **277**, 333–338 (1997).
4. Feig, L. *Curr. Opin. Cell Biol.* **6**, 204–211 (1994).
5. Lenzen, C., Cool, R., Prinz, H., Kuhlmann, J. & Wittinghofer, A. *Biochemistry* **37**, 7420–7430 (1998).
6. Fam, N. P. *et al.* *Mol. Cell Biol.* **17**, 1396–1406 (1997).
7. Kawashima, T. *et al.* *Nature* **379**, 511–518 (1996).
8. Wang, Y., Jiang, Y., Meiering-Voss, M., Sprinzl, M. & Sigler, P. B. *Nature Struct. Biol.* **4**, 650–656 (1997).
9. Renault, L. *et al.* *Nature* **392**, 97–101 (1998).
10. Cherfils, J. *et al.* *Nature* **392**, 101–105 (1998).
11. Geyer, M. *et al.* *Biochemistry* **35**, 10308–10320 (1996).

Earthquakes

Cracking Los Angeles

John H. Shaw

Los Angeles is caught in a vice, locked between the converging tectonic plates of North America and the Pacific. As a result, the Earth's crust beneath the city is shattered like glass, cut by faults of many shapes and sizes. Defining the motions along these faults, and the forces that may be released in earthquakes, is a complicated process, even when aided by satellites of the Global Positioning System (GPS). On page 356 of this issue¹, Walls *et al.* provide the latest estimates of these motions and thus of seismic risk. They conclude that a certain class of fault beneath Los Angeles, known as strike-slip and akin to the great San Andreas fault, may be more dangerous than has been thought.

The faults most likely to generate damaging earthquakes are the largest ones that show evidence of recent motions. Geologists study these faults by digging shallow trenches across them and boring into them to find

evidence of past earthquakes. If the past is a prologue, then these palaeoseismic records provide a way of predicting earthquake activity. Moreover, significant advances have been made in our ability to track fault movement using precise measurements of the GPS. These subtle motions 'load' (place stress on) faults and presumably give indications of future earthquakes.

Unfortunately, earthquake predictions based solely on either GPS or geological measurements don't always agree. Walls *et al.*¹ attempt to explain this discrepancy in the Los Angeles area by ascribing high rates of motion to strike-slip faults. This type of fault contrasts with thrust and reverse faults (Fig. 1), which have caused the most recent, large earthquakes in the Los Angeles area.

Three years of measurements from the Southern California Integrated GPS Network (SCIGN)² — a dense grid of fixed surface stations, the positions of which are

measured periodically using global positioning satellites — record about 8 mm yr⁻¹ of north–south shortening (contraction) across the Los Angeles basin. Palaeoseismic evidence presented by Walls *et al.*, and also by Rubin *et al.*³ in last week's *Science*, suggests that the Sierra Madre fault zone slips at a rate of about 1 mm yr⁻¹ (see Fig. 1 of the paper on page 357; the zone is part of a large reverse fault system that forms the San Gabriel Mountains). This modest slip rate, combined with rates predicted on other recognized fault systems, cannot fully account for the geodetically measured shortening. Moreover, Walls *et al.* report geodetically measured rates of east–west elongation, the converse of shortening, in Los Angeles that also exceed geologically determined values.

To explain these discrepancies, the authors ascribe higher rates of slip (up to 3 mm yr⁻¹) to conjugate strike-slip faults in the basin including the Verdugo, Raymond, San Jose and Chino. The authors suggest that these fault motions accommodate shortening and elongation by east–west block extrusion (Fig. 1a), which is analogous to the mechanism proposed to explain major strike-slip faults in China within and adjacent to the Tibetan Plateau⁴. Although poorly constrained, previous slip-rate estimates on these faults in the Los Angeles basin were generally half to one-quarter of the newly proposed values. If the proposed higher slip rates are correct, the chances of damaging earthquakes on these faults are higher than previously supposed⁵.

Historically, there have been no large earthquakes on the strike-slip faults to which Walls *et al.* ascribe the higher rates of slip. But, worldwide, the repeat times for earthquakes on faults appear to range widely, from decades to perhaps thousands of years. So that lack of large strike-slip events neither supports nor weakens the authors' predictions of higher slip rates. The relatively small size of the individual strike-slip faults suggests that earthquakes on them would be moderate in size and temporally clustered; the historical record may represent a quiescent period.

Alternatively, part of the geodetic and geological discrepancies could be explained by unknown or unappreciated fault systems, including 'blind' thrust faults^{6–8} which lie concealed beneath the Earth's surface (Fig. 1b). In 1994, the magnitude-6.7 Northridge earthquake⁹ ruptured a previously unrecognized blind thrust fault beneath a Los Angeles suburb, causing more than \$35 billion of damage. In fact, most recent large earthquakes beneath Los Angeles have been on thrust and reverse faults, not strike-slip faults, as Walls *et al.* acknowledge. Based on this record and the incomplete inventory of blind thrusts, as illustrated by the surprising Northridge earthquake, blind thrust and

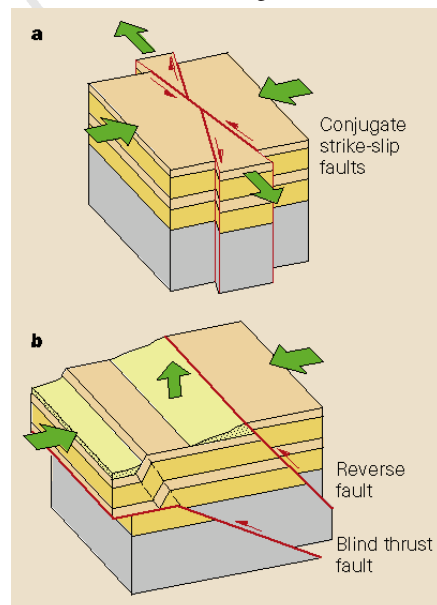


Figure 1 Perspective views of the two classes of faults that cause damaging earthquakes in the Los Angeles area. a, Strike-slip faults are generally vertical fractures across which blocks of Earth move laterally with respect to one another. Two or more faults that form at prescribed angles to one another are said to be conjugate. Conjugate strike-slip faults accommodate shortening in one direction and elongation in an orthogonal direction (denoted by green arrows). b, Thrust and reverse faults are inclined fractures where the blocks above the faults move up with respect to the underlying blocks. Thrust faults dip less than 45°, reverse faults dip more. Both thrust and reverse faults can be 'blind', meaning that they do not reach the Earth's surface. Yet blind faults may cause near-surface folding that defines their whereabouts. Thrust and reverse faults accommodate shortening in one direction and vertical uplift (denoted by green arrows).

reverse faults must still be considered both as significant hazards and partial sources of the observed geodetic and geological discrepancy.

Fortunately, these outstanding problems are tractable. Continued geodetic measurements using the SCIGN array will refine and reduce errors in estimates of fault motion. Palaeoseismic investigations may show whether the critical strike-slip faults move at the proposed high rates and whether past earthquakes have been clustered in time. Moreover, a wealth of new subsurface data recently made available by the petroleum industry is helping to define additional blind thrust faults and sources of hazard^{8,10}. These advances in characterizing seismic sources are complemented by better methods of forecasting ground shaking and of planning emergency responses. It seems inevitable that, sooner or later, a large earthquake will

occur in the Los Angeles area. Taken together, however, all of these approaches offer ways of mitigating the consequences. □

John H. Shaw is in the Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138, USA. e-mail: shaw@eps.harvard.edu

1. Walls, C. et al. *Nature* **394**, 356–360 (1998).
2. Bock, Y. et al. *J. Geophys. Res.* **102**, 18013–18033 (1997).
3. Rubin, C. M., Lindvall, S. C. & Rockwell, T. K. *Science* **281**, 398–401 (1998).
4. Tapponnier, P. P., Peltzer, G., Le Bail, A. Y., Armijo, R. & Cobbold, P. *Geology* **10**, 611–616 (1982).
5. Dolan, J. et al. *Science* **267**, 199–205 (1995).
6. Stein, R. & Yeats, R. F. *Sci. Am.* **260** (6), 48–57 (1989).
7. Davis, T. L., Namson, J. & Yerkes, R. F. *J. Geophys. Res.* **94**, 9644–9664 (1989).
8. Shaw, J. H. & Suppe, J. *J. Geophys. Res.* **101/B4**, 8623–8642 (1996).
9. US Geologic Survey and SCEC Scientists *Science* **266**, 389–397 (1994).
10. Schneider, C. L., Hummon, C., Yeats, R. S. & Huftile, G. L. *Tectonics* **15**, 341–355 (1996).

Palaeontology

An Asian Grande Coupure

Jean-Louis Hartenberger

The more palaeontologists know about mammals from the Cenozoic era (65 million years, Myr BP, to present), the more convinced they are that these fauna have undergone two major turning points in the Northern Hemisphere. The first is marked by the dispersal and subsequent radiation of most modern orders of mammals, within continents, at the Palaeocene/Eocene boundary, 55 Myr BP. The second happened 20 Myr later, around 34 Myr BP, at the Eocene/Oligocene boundary (EOB). This second event has been documented on a much more global scale than the first and, on page 364 of this issue, Meng and McKenna¹ add a new cornerstone for understanding it. They give an overview of successive mammalian faunas in Mongolia, from the Palaeocene to the end of the Oligocene, and identify a large-scale extinction — the Mongolian remodelling — which occurred at roughly the same time as the EOB.

The story starts at the beginning of the century when Hans Georg Stehlin², a Swiss palaeontologist, tried to convince his colleagues at the Société Géologique de France that the EOB, as observed in the marine realm, could also be recognized on land. From his studies of European Palaeogene faunas, he showed that the large herbivores suffered a big break at about the same time as the EOB. This event — which he named the *Grande Coupure* (the big cut) — corresponds to an extinction/origination event whereby new immigrants of Asian origin replaced most of the (old) endemic Eocene genera of Europe.

In 1966, Louis Thaler³ emphasized the point by showing that, at around the time

of the EOB, rodent biotas in Europe were enriched by Asian forms. A few years later, I proposed⁴ that the shift from a warm to a cooler climate at the end of the Eocene must have been a factor in this European event. Then, as debates about mass extinctions expanded after Alvarez and colleagues⁵ published their paper on the iridium and mass-extinction events at the Cretaceous/Tertiary boundary, those who studied the EOB decided to ask exactly what happened at the end of the Eocene. Most of them thought — and still think — that the boundary corresponds more to a transition than to a single catastrophic event. Nevertheless, at a symposium⁶ held in 1985 it was proposed that the EOB should be labelled the 'terminal Eocene event' (TEE), a name suggestive of a single event.

On a global scale, the TEE is marked by a climatic crisis that probably started with glaciation in the Antarctic Ocean and led to climate change in the Northern Hemisphere. In 1989, North American mammalian faunas were recalibrated and it was shown⁷ that important changes also occurred in this area, with a stepwise pattern near the EOB. So, Central Asian invaders were thought to have played a big part in the reworking of European and North American faunas, allowing many Asian families and genera to reach a cosmopolitan status.

With the synthesis of Meng and McKenna¹, which embraces a review of the available fossil record and recalibration of Mongolian formations, we can now appreciate what consequences the misnamed TEE had for the evolution of Central Asian faunas. The

authors find that after the many sizes of mammal that existed in the warm Eocene, small rodents and lagomorphs (hares and rabbits) predominated in the Oligocene. These mammals were more adapted to the arid and cool environment of the Oligocene and, in many ways, these changes were very like the European ones⁸.

Many questions remain. For example, how quickly did these species evolve, and what units of evolution are most influenced by climatic change? In Europe, there was a considerable change in the structuring of mammalian communities at the time of the TEE, so the *Grande Coupure* is not only a consequence of an Asian invasion⁸. In North America, there is much debate as to whether the climatic changes of the TEE have any influence on how long mammalian species existed, and whether evolutionary changes showed a bimodal pattern — long periods of stability interrupted by rapid evolutionary restructuring^{9,10}.

In my opinion, the most interesting question about the EOB lies in Africa and concerns our own family, Anthropeidae. The best fossil record that we have of its early radiation comes from the Palaeogene Fayoum deposits in Egypt¹¹. But what is the precise age of this rock formation? Were our early relatives born before or after the TEE? And did the climatic events of the TEE affect our origins? Because the position of the EOB is controversial in the Fayoum deposits^{12,13}, the best way to resolve this debate would be to look at other places in the southern margin of Tethys (which includes all of the countries from Morocco to Egypt, and part of Arabia). To work out when and why our very old ancestors emerged, we need more information on the Eocene-to-Oligocene history of mammals in Africa — only then should we know if the TEE had some influence on our evolution. □

Jean-Louis Hartenberger is in the Laboratoire de Paléontologie, Institut des Sciences de l'Évolution, Université de Montpellier II, place Eugène-Bataillon, 34095 Montpellier cedex 5, France. e-mail: hartenjl@isem.univ-montp2.fr

1. Meng, J. & McKenna, M. C. *Nature* **394**, 364–367 (1998).
2. Stehlin, H. G. *Bull. Soc. Géol. Fr.* **9**, 488–520 (1909).
3. Thaler, L. *Mém. Mus. Hist. Nat. 3^e sér. Sciences de la Terre* **17**, 1–295 (1966).
4. Hartenberger, J.-L. *Bull. Mus. Hist. Nat. Paris, 3^e sér. Sciences de la Terre* **24**, 49–70 (1973).
5. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
6. Pomeroy, C. & Premoli-Silva, I. R. (eds) *Terminal Eocene Event* (Elsevier, Amsterdam, 1986).
7. Prothero, D. R. in *Mass Extinctions: Processes and Evidence* (ed. Donovan, S. K.) (Columbia Univ. Press, New York, 1989).
8. Legendre, S. *Münchener. Geowiss. Abh.* **16**, 1–110 (1989).
9. Prothero, D. R. & Heaton, T. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 257–283 (1996).
10. Alroy, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 285–311 (1996).
11. Gagnon, M. J. *Hum. Evol.* **32**, 133–160 (1997).
12. Kappelman, J., Simons, E. L. & Swisher, C. C. I. *J. Hum. Evol.* **22**, 495–503 (1992).
13. Gingerich, P. D. *J. Hum. Evol.* **24**, 207–218 (1993).

Pulsars

A period of change

Nicholas E. White

On pages 344 and 346 of this issue, Wijnands and van der Klis¹, and Chakrabarty and Morgan², report the discovery of the first accreting millisecond X-ray pulsar. This bridges a vital gap in our theory of how pulsars evolve. It should also help us to test the predictions of General Relativity, and directly measure some of the properties of super-dense nuclear matter.

Neutron stars are the remnants of supernova explosions, packing a mass as great as that of the Sun into a radius of only about 10 km. Some, the millisecond pulsars³, also rotate at a breathtaking rate, up to about 1,000 revolutions per second. First discovered in 1982 from their radio emissions, these are the most rapidly rotating stars known. Unlike the more slowly spinning ordinary radio pulsars, which are newly born and slowing down, millisecond pulsars are believed to be very old neutron stars that have been recycled (spun up) by material consumed from a companion star, carrying angular momentum to the pulsar (Fig. 1, and <http://universe.gsfc.nasa.gov/videos/millisecond.html>). During this recycling phase, the neutron star should appear as a very bright X-ray source, because of the gravitational energy released by the accreted material. But in spite of this solid prediction, searches over the past 15 years had not, until now, turned up an accreting millisecond pulsar that emits X-rays.

There are, instead, many bright X-ray sources that do not show large-amplitude, coherent pulsations. Most of them are old X-ray binaries, where the companion star has a long life (like our Sun). Although some may

harbour black holes, most are thought to contain the progenitors of millisecond radio pulsars. If these are old neutron stars with magnetic fields that have decayed by a factor of 1,000 or more from their birth values, then the neutron star may be buried deep in a thick disk of accreting matter (a larger magnetic field would sweep the inner regions of the disk clear). The surrounding material may be disguising buried X-ray pulsars, by 'smearing out' the direct pulse from the pulsar, giving it a very low-amplitude modulation.

A new satellite called the Rossi X-ray Timing Explorer⁴ (RXTE) has now solved this problem. RXTE, part of NASA's Structure and Evolution of the Universe Program, is the largest collecting area X-ray telescope yet built. Because of its ability to catch so many photons, it can search for subtle timing effects, on timescales as short as microseconds, in systems containing neutron stars and black holes in our Galaxy. Observations with RXTE over the past year-and-a-half have turned up a wealth of new phenomena. The presence of neutron stars with millisecond periods has already been inferred from the discovery by RXTE of quasi-periodic oscillations (QPOs) with frequencies⁵ of 1,000 Hz, and from 300-Hz oscillations seen during short-lived X-ray bursts⁶ from other sources. The QPOs are believed to come from the beat between the neutron-star rotation period and the inner orbital period of the accretion disk.

But these earlier results, although extremely suggestive⁷, are not proof of a millisecond neutron-star rotation period. The 2.5-ms period of the new X-ray pulsar, called

J1808–369, at last provides that proof^{1,2}. Millisecond radio pulsars are indeed born in X-ray binaries. This confirms a key element of our picture of pulsar evolution.

This is a new type of pulsar. Most accreting X-ray pulsars have much longer periods, of seconds to hundreds of seconds, and are typically in young systems where their companion is a massive, short-lived O- or B-type star. The magnetic fields of these young pulsars are huge — around 10^{12} gauss, one trillion times as strong as the Earth's field — creating a large magnetosphere around the neutron star. The pulsar is spun up until its rotation period equals the orbital period of gas at the accretion-disk-magnetosphere boundary. (If it were to go any faster, the rigid magnetic field would drag against the accretion disk, slowing the neutron star down again.) Hence, a millisecond X-ray pulsar requires a neutron star with a much weaker magnetic field (weaker by a factor of 1,000), so that the size of the magnetosphere is only a few tens of kilometres, corresponding to an orbital period of a millisecond.

It is surprising that so few X-ray binary systems appear to have a millisecond pulsar. An important clue to the reason for this may be the fact that this first example is a transient source, visible only when there is a sporadically increased flow of material from the companion star. If all millisecond X-ray pulsars are transients, we will have missed most of them in ordinary surveys because they will have been off.

In this pulsar and other X-ray binary systems, we see emission from material close to or at the neutron star surface. This may enable us to constrain the equation of state of nuclear matter by constraining the mass and radius of the neutron star. It also is a regime where the effects of General Relativity are obvious, and so it should allow a direct test of the theory's fundamental predictions. For example, RXTE has already delivered evidence for the marginally stable orbit predicted by General Relativity⁸. This new pulsar will help us make further progress, because it is a very stable clock in the deep potential well of a neutron star. Already, the pulse timing measurements presented here have precisely determined the two-hour orbit of one star about the other. And if this transient source stays active for another year, we will be able to put limits on the decay of the companion-star orbit², and determine whether it is caused by gravitational radiation or other effects such as magnetic braking. □

Nicholas E. White is in the Astrophysics Branch, NASA Goddard Space Flight Center, Code 662, Greenbelt, Maryland 20771, USA.

e-mail: nwhite@heapop.gsfc.nasa.gov

1. Wijnands, R. & van der Klis, M. *Nature* **394**, 344–346 (1998).
2. Chakrabarty, D. & Morgan, E. H. *Nature* **394**, 346–348 (1998).
3. Backer, D. C. et al. *Nature* **300**, 615–618 (1982).
4. <http://rxte.gsfc.nasa.gov>
5. van der Klis, M. et al. *Astrophys. J. Lett.* **469**, L1–L4 (1996).
6. Strohmeyer, T. E. et al. *Astrophys. Lett.* **469**, L9–L12 (1996).
7. White, N. E. & Zhang, W. *Astrophys. J. Lett.* **490**, L87–L90 (1997).
8. Zhang, W. et al. *Astrophys. J. Lett.* **500**, L171–L174 (1998).



Figure 1 A star is reborn: artist's impression of the J1808–369 system. Gas flows from a star that is about 10–20% the size and mass of the Sun onto its tiny but more massive companion, a neutron star. This material forms an accretion disk around the neutron star, spinning it up until the star's magnetosphere and the accretion disk's inner edge are co-rotating. X-rays come from matter hitting the neutron star's surface, and from thermonuclear flashes in matter that builds up there. When no more material can be stripped from its companion, the neutron star will become a millisecond radio pulsar. The two stars are about as far apart as the Earth and the Moon.

Molecular motors

Keeping up with the F_1 -ATPase

Howard C. Berg

Well, seeing is believing. Reporting in *Cell*, Kinosita, Yoshida and colleagues¹ have once again turned molecules of F_1 -ATPase on their heads, visualizing rotation of the γ -subunit relative to the three alternating sets of α - and β -subunits. As they did last year², the authors have attached a histidine tag to one end of the β -subunits and inverted them over glass (although now elevated by a nickel-coated bead, 0.2 μm in diameter), and attached an actin filament to the γ -stalk through biotin-streptavidin-biotin links (Fig. 1). This time, however, they have used much lower concentrations of ATP to show that F_1 rotates the actin tag in discrete 120° steps at exponentially distributed intervals. Not only that, but the work done by each step is close to the energy available from the hydrolysis of one molecule of ATP, and this is true for actin

filaments of different lengths (over a range of viscous loads).

Is the near-100% efficiency with which the F_1 -ATPase converts chemical to mechanical energy surprising? Kinosita and colleagues^{1,3} think not, because this accords with the fully reversible nature of the native enzyme, which is composed of two parts, F_1 and F_0 (Fig. 2a). When given ATP, F_1 drives F_0 , which pumps protons, establishing an electrochemical gradient across the cytoplasmic membrane (inside negative). When provided with protons energized by an electrochemical gradient, F_0 drives F_1 , which synthesizes ATP.

The only other proton-driven rotary device is the flagellar motor (Fig. 2b). Available evidence⁴, although not as good as that obtained by Kinosita and colleagues, indicates that this also runs at high efficiency —

but only when turning slowly. Molecular motors are not heat engines. If the energy available from the hydrolysis of ATP were uniformly distributed throughout the F_1 -ATPase, it would heat up by less than 0.1 °C and then cool off with a decay time of less than 0.1 ns. The energy available from the binding of ATP or the release of inorganic phosphate (or whatever) must be stored in springs (conformational or electrostatic) which, on relaxing, drive the actin filament through 120°. The details of this mechanism should prove fascinating.

How does the F_0 part of the ATPase work? The main problem is that we are not yet certain which components comprise the stator and which the rotor. Most people believe that protons move between the a and c subunits, interacting with the Arg 210 residue in a and the Asp 61 residue in c (Fig. 2a). This view has been developed quantitatively by Elston *et al.*⁵, in a version of a thermal-ratchet model that was invented earlier for the flagellar motor⁶. The proton rides with Asp 61 on the rotor. The pK_a of Asp 61 is then lowered by interaction with Arg 210, forcing the proton off whenever the rotor, moving thermally, tries to go the wrong way. One proton is carried per c subunit, 12 per revolution, four per ATP synthesized. The Cys 205 residue of the γ -subunit has been crosslinked to c subunits at positions 42, 43 or 44 without loss of ATPase activity⁷, but it is difficult to tell whether this alters proton pumping. The interacting helical faces of a and c have also been defined by crosslinking⁸, although it is not known whether such crosslinks block function.

We are not much better off with the flagellar motor (Fig. 2b), although we know that this motor turns a propeller relative to the rigid framework of the cell. So, the torque generators must be bolted down to the peptidoglycan somewhere, and the outer part of the rotor, drive shaft, flexible coupling and propeller must turn as a unit. Again, the view presented in Fig. 2 is that of the majority — it is possible that the inner part of the rotor (the FliM or FliN components of the C ring) does not rotate.

It is tempting to compare the F_1 actin-filament assays^{1,2} with the tethered-cell assay⁹ in which the propeller is fixed to a glass slide and one follows the rotation of the cell body. But this experiment would be analogous to one in which the γ -subunit was fixed to a glass slide, and the rotation of a and b observed. Rotation of the γ -subunit relative to $\alpha_3\beta_3$ drives the synthesis of ATP, whereas the hydrolysis of ATP turns γ . For the flagellar motor, rotation of the flagellar filament moves fluid (enabling the cell to swim) whereas imposing flow on the filament causes it to rotate. (The latter experiment was done nearly 20 years ago by Hotani¹⁰.) In both systems, more structural information

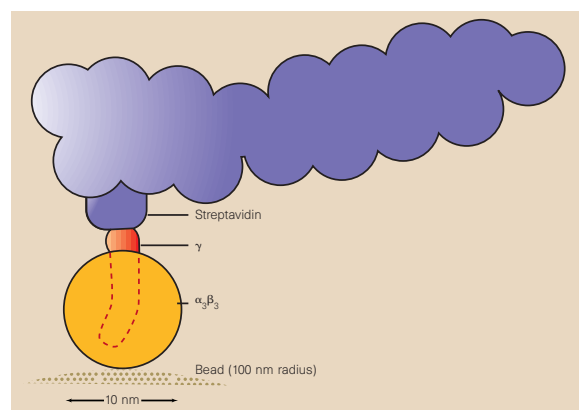


Figure 1 The experimental set-up of Kinosita and colleagues¹, drawn to scale. The α - and β -subunits of F_1 are linked to a bead sitting on a glass slide, and an actin filament is linked to the γ -subunit. When ATP is added, the actin filament rotates, showing that ATP hydrolysis drives rotation of the γ -subunit relative to $\alpha_3\beta_3$. The actin filaments used in the experiments were 10 to 100 times as long as the segment shown here.

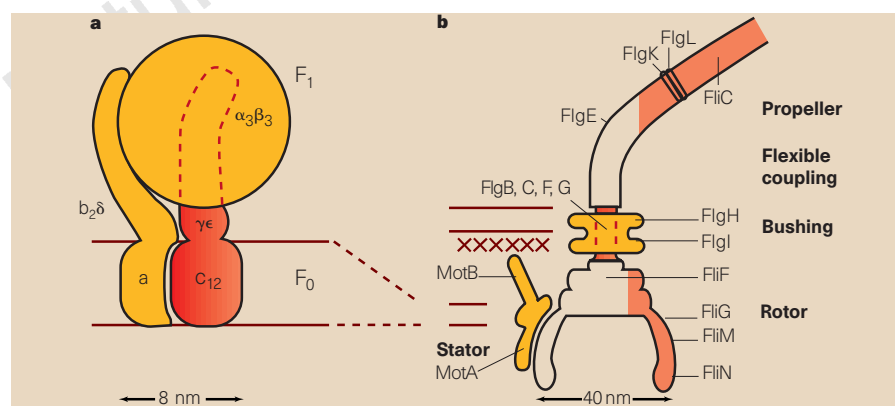


Figure 2 Two rotary machines of *Escherichia coli*. a, F_1F_0 -ATPase, shown cytoplasmic-side up. The components $a\delta b_2\alpha_3\beta_3$ are thought to serve as the stator (yellow), $\gamma\epsilon c_{12}$ as the rotor (red). The F_1 portion can be detached from the cytoplasmic membrane as a unit (components denoted in Greek letters); the F_0 portion remains embedded in the membrane (components denoted in Roman). b, Bacterial flagellum, shown cytoplasmic-side down. There are at least eight torque-generating stator elements, each linked to the peptidoglycan (the rigid framework of the cell wall, shown as X's) and containing the proteins MotA and MotB (yellow). The bushing (also yellow) gets the drive shaft through the outer layers of the cell wall (the peptidoglycan layer and the outer membrane, shown as a pair of horizontal lines). Both machines are drawn to scale, but the flagellum is reduced in size by a factor of five. (For a review of the earlier F_1 work², see Block¹⁵. For other methods used to show rotation in F_1 , see Junge *et al.*¹⁶.)

is available for external components^{11,12} than for those embedded in the membrane. But whereas the ATPase community can disassemble and reassemble their machine at will, work *in vitro* with the flagellar motor is more difficult.

A point to note is that Na⁺-driven versions exist for both of these machines, and the structures of many of the components in either version are homologous. An Na⁺-translocating ATPase of the anaerobic bacterium *Propionigenium modestum* has been extensively studied¹³, as have the flagellar motors of certain alkalophilic and marine bacteria such as *Vibrio alginolyticus*¹⁴. So one should think of these devices as ion driven, not proton driven.

Is ATP synthesis the only assay for F₀ output, or will other actin filaments be spinning in the future? □

Howard C. Berg is in the Department of Molecular and Cellular Biology and the Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA, and is a member of the Rowland Institute for Science, 100 Edwin H. Land Boulevard, Cambridge, Massachusetts 02142, USA. e-mail: hberg@biosun.harvard.edu

1. Yasuda, R., Noji, H., Kinosita, K. J. & Yoshida, M. *Cell* **93**, 1117–1124 (1998).
2. Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. J. *Nature* **386**, 299–302 (1997).
3. Kinosita, K. J., Yasuda, R., Noji, H., Ishiwata, S. & Yoshida, S. *Cell* **93**, 21–24 (1998).

4. Meister, M. & Berg, H. C. *Biophys. J.* **52**, 413–419 (1987).
5. Elston, T., Wang, H. & Oster, G. *Nature* **391**, 510–513 (1998).
6. Berg, H. C. & Khan, S. in *Mobility and Recognition in Cell Biology* (eds Sund, H. & Veeger, C.) 485–497 (de Gruyter, Berlin, 1983).
7. Watts, S. D., Tang, C. & Capaldi, R. A. *J. Biol. Chem.* **271**, 28341–28347 (1996).
8. Jiang, W. & Fillingame, R. H. *Proc. Natl Acad. Sci. USA* **95**, 6607–6612 (1998).
9. Silverman, M. & Simon, M. *Nature* **249**, 73–74 (1974).

10. Hotani, H. *J. Mol. Biol.* **129**, 305–318 (1979).
11. Abrahams, J. P., Leslie, A. G. W., Lutter, R. & Walker, J. E. *Nature* **370**, 621–628 (1994).
12. Namba, K. & Vonderviszt, F. *Q. Rev. Biophys.* **30**, 1–65 (1997).
13. Dimroth, P. *Biochim. Biophys. Acta* **1318**, 11–51 (1997).
14. Asai, Y. *et al. J. Bacteriol.* **179**, 5104–5110 (1997).
15. Block, S. M. *Nature* **386**, 217–219 (1997).
16. Junge, W., Lill, H. & Engelbrecht, S. *Trends Biochem. Sci.* **22**, 420–423 (1997).

Combinatorial chemistry

Connecting with catalysis

Ian E. Maxwell

During the past ten years, combinatorial chemistry has been developed and applied by the pharmaceutical industry in particular. This technology, which involves the micro-scale synthesis and screening of vast libraries of organic compounds, has resulted in great improvements in the efficiency of testing new molecules for drug activity. But only recently have the possibilities of using similar techniques in other areas been explored. One such example — high-speed screening for new solid-state catalysts — comes in the paper by Senkan on page 350 of this issue¹.

The history of research into 'other applications' is short, and stems from pioneering work carried out by Xiang, Schultz and

colleagues² at the University of California, Berkeley, in 1995. They showed that the approach could be used to explore, for example, complex metal-oxide compositions for potential use in superconductors. The commercial prospects can be gauged from the fact that Schultz became the co-founder of a new high-tech company, Symyx, whose core purpose is to employ combinatorial and high-throughput screening methods to search for new materials, including catalyst systems.

Senkan now describes a high-speed screening technique for searching for new heterogeneous catalyst systems. The experimental set-up is depicted in Figs 1 and 2 of the paper (page 351). Senkan has con-

Palaeoanthropology

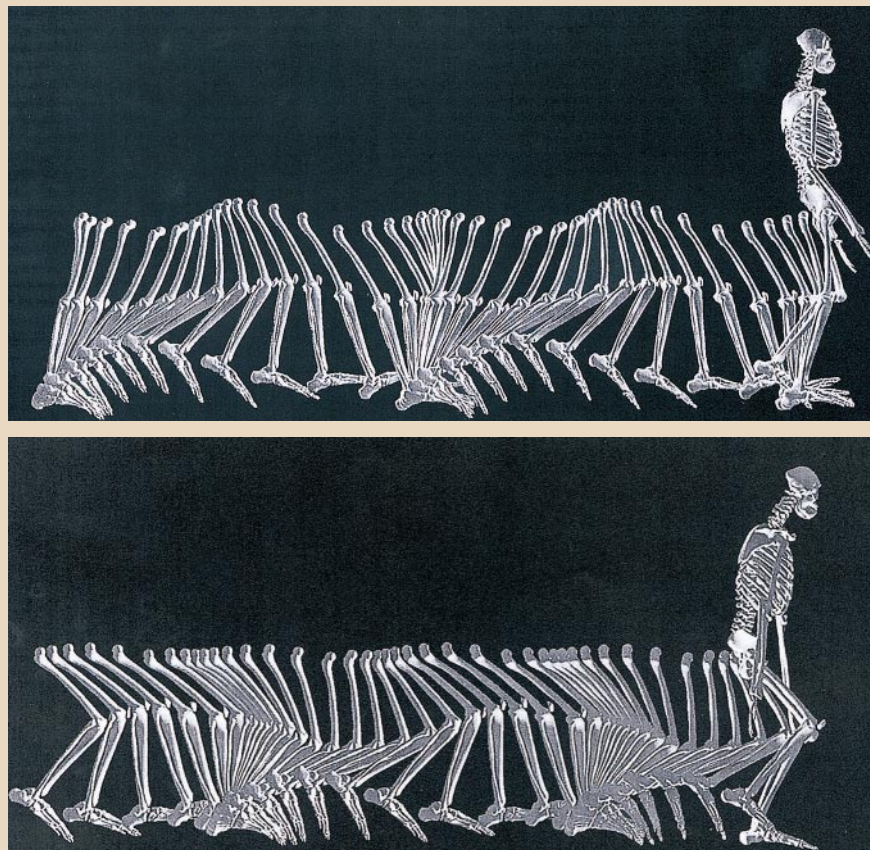
Lucy takes a stroll

The skeleton of Lucy, the best-known example of the hominid *Australopithecus afarensis*, shows that its owner probably got around on two legs. But it is not at all clear what gait was preferred. Was it chimpanzee-like? Or the related 'bent-hip, bent-knee' carriage that humans can use but normally don't (as seen in exaggerated form in films featuring Groucho Marx)? Or erect, like that of modern humans?

Robin Crompton and colleagues have tackled these questions by assessing the biomechanical effectiveness of a numerically simulated model skeleton of Lucy, and testing it under various assumptions (*Journal of Human Evolution* **35**, 55–74; 1998). Among the results are the screen dumps shown here, which come from animations run by the authors. They show Lucy with 'motion functions' derived from erect human walking (top) and bent-hip, bent-knee walking.

Crompton *et al.* rule out a chimpanzee-like gait. They also conclude that bent-hip, bent-knee motion is so much less mechanically effective, and heat generation so much greater, that an erect carriage was favoured. Lucy, they believe, walked tall.

Tim Lincoln



structed a 72-site (8×9) solid-state library of different catalyst formulations and screened for the catalytic dehydrogenation of cyclohexene to benzene. Although this particular catalytic reaction is not widely applied industrially, many other commercial processes — catalytic cracking, isomerization, oxidation, hydrogenation and even dehydrogenation — employ heterogeneous catalysts under similar gas-phase conditions.

Miniaturization is a key feature of these new combinatorial technologies, because it enables huge libraries of materials to be synthesized and screened at high speed and low cost with minimal space requirements. Senkan has achieved miniaturization to the level of individual library sites with dimensions of about 5 mm by 5 mm, corresponding to a library density of about ten sites per 5 cm^2 . The analytical technique involves the selective photoionization of reaction products, using tuneable ultra-violet lasers under conditions known as resonance-enhanced multiphoton ionization (REMPI).

Other developments in this area include the demonstration³ that combinatorial techniques can be applied to the synthesis of zeolites, which are the key components of many important catalyst systems. Another impressive example⁴ is the use of a combinatorial library to screen for electrocatalysts for the methanol–air fuel cells that could potentially power vehicles having ultra-low emissions of pollutants.

In this second case, the group concerned deployed ink-jet technology to create an array of 645 catalyst ‘dots’, each with a different composition. The array was screened for electrocatalytic activity using a fluorescent acid–base indicator which emits visible light when hydrogen ions (a product from methanol decomposition) are produced. Using this tool, the group discovered a new quaternary catalyst for methanol fuel-cell anodes based on a combination of the metals platinum, ruthenium, osmium and iridium. It produces a current density 30–60% higher than a commercial catalyst that uses a binary platinum–ruthenium system.

As catalytic systems become more complex, the power of these new approaches becomes even greater. That could be of immense economic importance. Use of catalysts is pervasive — in oil refining, in producing petrochemicals, fine chemicals and polymers, and in power generation, for example (put another way, catalytic applications indirectly contribute to some 20–30% of the gross domestic product of developed countries). Furthermore, the search for new catalysts tends to be haphazard because of our limited scientific understanding of how they work. So combinatorial and high-throughput screening methods have attracted a great deal of

interest from the academic community, as well as industry.

It is difficult to predict future developments in a rapidly emerging technology such as this, not least which applications will enjoy early success (promising areas include low-pressure, gas-phase heterogeneous catalysis, polymer catalysis, electrocatalysis, electronic materials and microporous materials). But it seems that sites 200 micrometres square will be feasible, which would result in about 10,000 different materials per 5 cm^2 — a great advance in screening rates compared with conventional technologies.

Senkan’s approach¹ elegantly demonstrates ‘proof of principle’ for screening heterogeneous catalyst systems, but it does have some limitations. For example, although the REMPI technique is suitable for detecting aromatic molecules, it may be not be appropriate for reactions involving other compounds such as olefins and paraffins. Another drawback is the mixing of product gas streams in the screening configuration chosen and the lack of mass balances. However, as Senkan says, these problems could be tackled by a smaller chamber design, higher gas-flow rates and the use of a honeycomb channel structure. Many of these problems are generic to gas-phase reaction systems in heterogeneous catalysis, and solutions will undoubtedly be found in the near future.

One of the biggest challenges, both here and in other areas, is likely to lie in the development of analytical techniques that realistically relate to material performance, rather than the library mechanics. Different approaches will probably be needed depending on the type of application. Such is the pace of development, however, that the ‘lab on a chip’ may not prove to be a dream.

Finally, one cause for optimism in the ‘other applications’ arena comes from a problem that is largely confined to pharmaceutical research. Here, combinatorial techniques have indeed dramatically cut the time necessary for early drug screening. But the lengthy period (on average about ten years) required for clinical trials has hardly shortened, and so the approach has had little impact on the overall time needed to bring new drugs to the market-place. This regulatory bottleneck does not exist for materials and catalyst systems — the effect on speed-to-market of combinatorial and high-speed screening could therefore be much greater. □

Ian E. Maxwell is at the Shell Research and Technology Center, PO Box 38000, 1030 BN Amsterdam, The Netherlands.
e-mail: Ian.E.Maxwell@opc.shell.com

1. Senkan, S. M. *Nature* **394**, 350–353 (1998).
2. Xiang, X.-D. et al. *Science* **268**, 1738–1740 (1995).
3. Akporiaye, D. E., Dahl, I. M., Karlsson, A. & Wendelbo, R. *Angew. Chem. Int. Edn Engl.* **37**, 629–631 (1998).
4. Reddington, E. et al. *Science* **280**, 1735–1737 (1998).

Daedalus

Wet metallic hydrogen

Many efforts have been made to crush hydrogen into a metallic state, so far without success. A metal is an array of positive metal ions, surrounded by a sea of mobile, delocalized electrons. In metallic hydrogen, if we could make it, the ions would presumably be simply protons.

Daedalus now points out that while hydrogen ions feature prominently in chemistry, they never, ever, exist as free protons. Always they attach themselves to some molecule or other. Thus in water, the proton “H⁺” exists as the hydroxonium ion H₃O⁺. So, he says, the search for metallic hydrogen is wrong-headed. The way forward is to exert pressure, not on pure hydrogen, but on a solution of hydrogen in water. The result would be a new metal, hydroxonium. It would be an array of H₃O⁺ ions in an electron sea. It should be far easier to make than the putative metallic hydrogen itself; with luck, it will form at quite low pressures.

Hydroxonium will be a sort of metallic ice. Once it has been compressed into existence, it will be resonance-stabilized by the delocalization of its conduction electrons. So it should stay stable if the pressure is then reduced. Daedalus even dreams that it might remain stable right down to normal pressures. Even if it does, hydroxonium will probably not be much use in engineering. Like sodium, with which it is isoelectronic, it will be too soft and reactive. Indeed, it may be so wildly reactive as to be a splendid rocket fuel. On heating, it should decompose to steam, hydrogen and a great deal of energy.

Hydroxonium might be better employed as a minor constituent of metal alloys. Even highly reactive metals (such as lithium) can form useful alloys. Diluted with stabler metals, they lose their reactivity while still conferring valuable properties, such as high tenacity and low density, on the resulting alloy.

But Daedalus reckons that nature may have got there first. Many comets, asteroids and satellites are taken to be ‘dirty snowballs’ — water ice with impurities. Hydrogen, the most abundant element in the cosmos, will be one such impurity. So the gravitationally compressed centre of the biggest dirty snowballs may consist of hydroxonium. Space visionaries have often proposed a self-supporting form of space exploration, in which each planetary body is mined for water and rocket fuel to reach the next. Hydroxonium could make it feasible.

David Jones

Wheatstone's waves

The nineteenth-century creative genius Sir Charles Wheatstone invented a wave machine and other 'philosophical toys' that had a serious purpose in demonstrating the laws of physics.

Martin Kemp

Gaining a visual grip on compound motions in more than one plane is a tricky business, particularly when the mechanism cannot be observed directly. Under these circumstances, the didactic machine or what was called the 'philosophical toy' comes into its own.

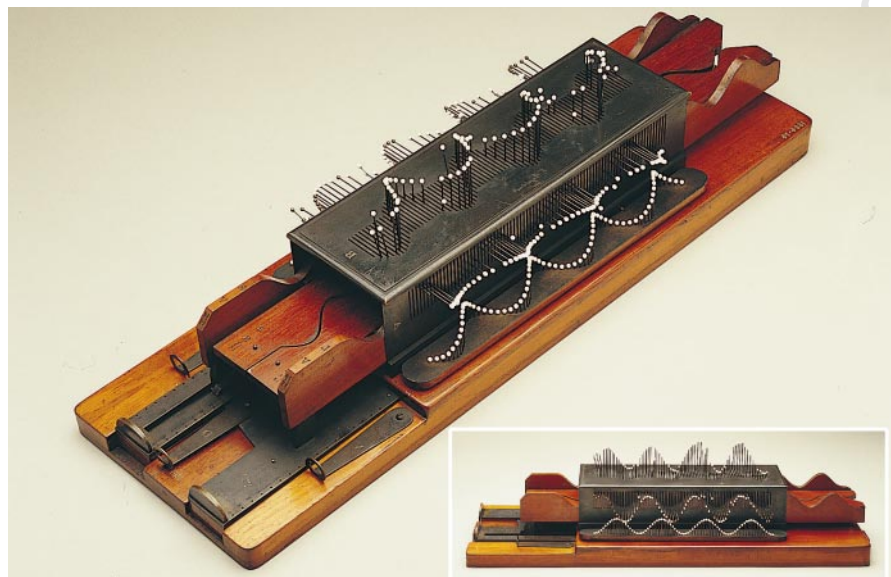
The most obvious role for such devices is the inducing of the 'Ah-hah!' factor – that instant of illumination when we find that we have suddenly acquired the elusive visualization of how a complicated system works in spatial terms.

In addition to instruction in known mechanisms, 'philosophical' machines could model hypotheses in a more speculative manner, acting to confirm the feasibility of a theory. Exemplary cases are the armillary sphere, which modelled the Ptolemaic system of the cosmos, and Wheatstone's "Wave Machine", which gave visible shape to the interference of different wave-forms.

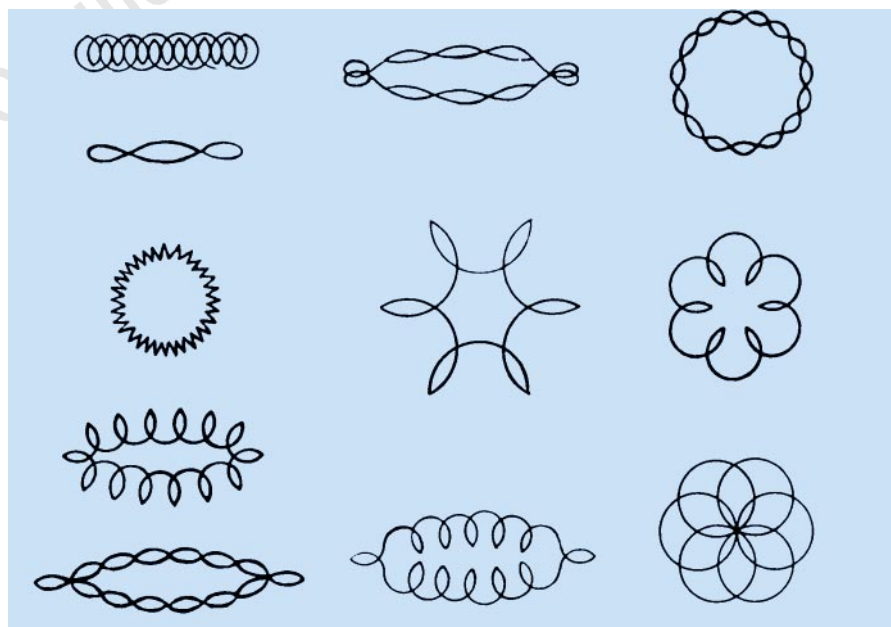
Best known as a pioneer of the electric telegraph, Sir Charles Wheatstone (1802–75) came to his main sciences of acoustics, optics and electricity via the making and inventing of musical instruments. The enduring theme of his creative life may be described as the invention of mechanisms of direct practical utility and didactic cunning that were profoundly in tune with natural law.

Wave motion, such a crucial theoretical model in physics from the time of Thomas Young and Augustin Fresnel in the early nineteenth century to the present day, was the subject of the most compelling of Wheatstone's didactic toys. His wave machine was invented in the early 1840s and produced for sale in handsomely engineered versions. It ingeniously represents the motion of 'ether particles' in plane, elliptical and circular polarization through the compounding of two sinusoidal motions at right angles.

White beads of mother-of-pearl are located at the tips of wiry rods that move sinusously in response to a set of more than two dozen wave-shaped, wooden templates inserted in pairs into the machine at right angles to one another. The instructions issued in 1884 by John Newman, "Philosophical Instrument Maker", give a nice idea of the kinds of shapes supplied: "The slides are marked R & L (right and left) H & V (Horizontal and Vertical); (1) The standard wave; (2) The lower octave, 2:1; (3) The major third below, 5:4."



Wheatstone's Wave Machine at the Science Museum in London.



Wheatstone's "Sonic Patterns Produced by a Kaleidophone".

A complementary realization of the 'musical' principles governing the 'molecular vibrations' in both light and sound was provided by Wheatstone's "Kaleidophone or Phonic Kaleidoscope... for the Illustration of Several Interesting and Amusing Acoustical and Optical Phenomena". Named after Sir David Brewster's optical kaleidoscope, it used small illuminated mirrors attached to the top of vibrating rods that emerged from a base plate in which varied vibrations were induced. The paths of the oscillating mirrors, traced by the persistent impressions of

reflected light, created beautiful and surprising patterns, comparable to the Chladni's figures produced by sand on vibrating plates.

For Wheatstone, as for Brewster, the philosophical toy was designed to induce simple delight in nature's concealed wonders, particularly in the young, and to model in mechanical form profound truths about physical law at the highest levels of understanding. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

DNA microsatellite analysis of Dolly

Dolly, the first animal cloned from an adult mammal, was produced by somatic cell nuclear transfer from a cell population derived from mammary tissue taken from a 6-year-old Finn Dorset ewe¹. Analysis of DNA from Dolly showed that she contained the same seven microsatellite alleles as those present in the cell population from which she was derived¹. Here we report a more detailed microsatellite analysis, which confirms the origin of Dolly.

Sgaramella and Zinder² recently asserted that sheep are "highly inbred", and they queried whether the microsatellite evidence that we originally provided¹ was sufficiently robust to exclude the possibility that Dolly was derived from embryonic or fetal cells from a different animal. They also suggested that, because the ewe was pregnant, Dolly might have been derived from a stray fetal cell contaminating the mammary tissue rather than from a mammary cell itself.

This hypothesis seems improbable as there are few fetal cells in the circulation of pregnant women (from 1 in 10⁵ to 1 in 10⁹ maternal cells³) and there is a much less intimate relationship between maternal and fetal circulation in sheep than in humans⁴. We think it highly unlikely that the partially purified mammary cells could have been overgrown by contaminating fetal cells during their relatively short time in culture⁵.

Sheep populations are usually outbred rather than inbred and the five microsatellites used in the previous analysis (data from four of which were published¹) were chosen because they are polymorphic in sheep⁶. However, in the absence of information about allele frequencies in the specific population from which the 6-year-old ewe was taken, we could not estimate how useful these markers are in discriminating between a fetal or maternal origin for Dolly. Thus we carried out the more detailed analysis described here.

The cells used to produce Dolly were prepared from mammary tissue taken from a 6-year-old ewe as part of a separate collaboration between the Hannah Research Institute in Scotland and PPL Therapeutics. A sample of the tissue that had been stored, frozen, at the Hannah Research Institute since the ewe was killed in 1995, and DNA from the original cell populations, was provided by C. Wilde for this analysis. We calculated the frequency of alleles in the Finn Dorset flock at the Hannah Research Institute from data from 44 individuals drawn from two separate generations but with only one representative from each full-sib family. Blood samples were delivered to Rosgen, a specialist genotyping company set up by the Roslin Institute and accredited

Table 1 Microsatellite analysis of Dolly

Marker	Number of alleles found in Finn Dorset population	Alleles present in Dolly, mammary tissue and cultured cells	
		Size (base pairs)	Frequency in population (standard error)
TGLA53	7	151/151	0.55 (0.05)
SPS 115	4	248/248	0.22 (0.05)
TGLA126	7	118/126	0.17 (0.04)/0.24 (0.05)
TGLA122	9	190/190	0.07 (0.03)
ETH3	4	104/106	0.21 (0.04)/0.49 (0.05)
ETH225	4	148/150	0.62 (0.05)/0.02 (0.02)
FCB11	6	124/126	0.13 (0.04)/0.14 (0.04)
MAF209	4	109/121	0.17 (0.04)/0.57 (0.05)
FCB128	5	112/112	0.49 (0.06)
ETH10	1	208/208	1.00 (0.00)

to the quality standard ISO9002. DNA was extracted using standard protocols.

Microsatellite amplification was done using three of the primer sets used previously and seven proprietary markers from Perkin Elmer, and the products of DNA amplification by the polymerase chain reaction were analysed using an ABI 377 Prism Sequencer. The markers from Perkin Elmer were originally developed for parentage testing in cattle but had also been shown to be polymorphic in sheep. The presence or absence of each allele was determined in DNA extracted from blood from Dolly and the other Finn Dorset animals, from cells from the original mammary tissue, and from cells at passage four (Dolly was derived from cells at passage three).

The data on allele frequencies are summarized in Table 1. All but one of the ten microsatellite markers were polymorphic, with from four to nine alleles present per marker. The alleles present in DNA from Dolly were identical to those in the original mammary tissue, in the cell population prepared from that tissue and in the cells cultured to passage four. We estimated the probability that another sheep from the same population would have the same genotype as the 6-year-old ewe to be between 1.9×10^{-12} and 2.7×10^{-10} (95%

confidence interval). We conclude that it is extraordinarily unlikely that Dolly was derived from a different Finn Dorset animal and, therefore, reject the hypothesis that "imagined and unimagined experimental error"² occurred.

If Dolly were derived from a fetal cell, she would have derived half of her alleles from the sire of the fetus and half from the 6-year-old ewe. We calculated the chance of a fetal cell having the same genotype as the 6-year-old ewe to be between 1.1×10^{-6} and 9.2×10^{-6} (95% confidence interval). We conclude that Dolly was derived from a mammary cell of the 6-year-old donor ewe.

David Ashworth, Matthew Bishop

Rosgen, Roslin, Midlothian EH25 9PS, UK

Keith Campbell, Alan Colman, Alex Kind, Angelika Schnieke

PPL Therapeutics Ltd, Roslin,

Midlothian EH25 9PP, UK

Sarah Blott, Harry Griffin, Chris Haley,

Jim McWhir, Ian Wilmut

Roslin Institute, Roslin, Midlothian EH25 9PS, UK

e-mail: ian.wilmut@bbsrc.ac.uk

1. Wilmut, I. *et al. Nature* **385**, 810–813 (1997).

2. Sgaramella, V. & Zinder, N. D. *Science* **279**, 635–636 (1998).

3. Goldberg, J. D. *Am. J. Hum. Genet.* **61**, 806–809 (1997).

4. McDonald, L. E. *Veterinary Endocrinology and Reproduction* 455–459 (Lea & Febiger, Philadelphia, 1980).

5. Campbell, K. H. S. *et al. Science* **279**, 636–637 (1998).

6. Buchanan, F. C. *et al. Mamm. Genet.* **4**, 258–264 (1993).

DNA fingerprinting Dolly

The birth of Dolly¹ has raised considerable interest and debate over the potential for cloning mammals, including humans. However, there are concerns about the authenticity of Dolly, and whether she could have been derived not from an adult donor mammary cell, but instead from a contaminating sheep cell culture or from a fetal cell present in the udder of the pregnant ewe donor². Microsatellite typing suggested but did not prove authenticity^{1,2}. We

have therefore carried out a DNA fingerprint analysis to determine the origin of the donor cell used in nuclear transfer, and have confirmed the authenticity of Dolly.

The cell culture and micromanipulations used to make Dolly were done at PPL Therapeutics and the Roslin Institute, Edinburgh, using mammary cells prepared at the Hannah Research Institute from tissue of a late-pregnant Finn Dorset ewe³. Immunocytochemical analysis showed the

cell population to be predominantly (>98%) of epithelial origin, and able to express genes encoding milk proteins under permissive conditions. When cultured with lactogenic hormones on reconstituted basement membrane⁴, the cells expressed messenger RNAs encoding α_{s1} -casein, β -casein and β -lactoglobulin. Serial passaging, as used for nuclear transfer¹, eliminated the cells' ability to express milk-protein genes.

We used DNA fingerprinting to compare cells from the original culture with mammary tissue from the donor ewe that had been stored frozen at the Hannah Institute and with a blood sample from Dolly provided by the Roslin Institute. Taking of the blood sample was witnessed.

As the sheep DNA fingerprinting systems described⁵ are not highly discriminating, we used a cocktail of four different probes, each known to be effective at detecting variable minisatellites in various species^{6,7} (our unpublished data). Southern blot hybridization with this cocktail (Fig. 1) showed that there is extensive variability in the DNA of Dolly and 12 control sheep, including a mother/offspring pair, from the Finn Dorset flock from which Dolly's donor was taken. We scored 58 different variable bands in total in these animals, with on average 17 such bands per sheep (range 13–22). Only 38% of the variable bands were shared between any two animals.

In contrast, the DNA fingerprints of the mammary tissue, the primary cell culture and Dolly were indistinguishable in terms of band number, position and relative intensity. Assuming a lack of linkage and association between different DNA fingerprint bands, as has been established for most bands in other mammalian multilocus analyses^{8–10}, the probability that a second unrelated sheep has, by chance, the same profile as the donor tissue can be conservatively estimated^{9,10} at 6×10^{-10} . We therefore reject the possibility that Dolly was derived from a contaminating cell culture.

To determine whether Dolly could have been derived from a fetal cell, and could therefore be an offspring, rather than a clone, of the donor, we used band-sharing

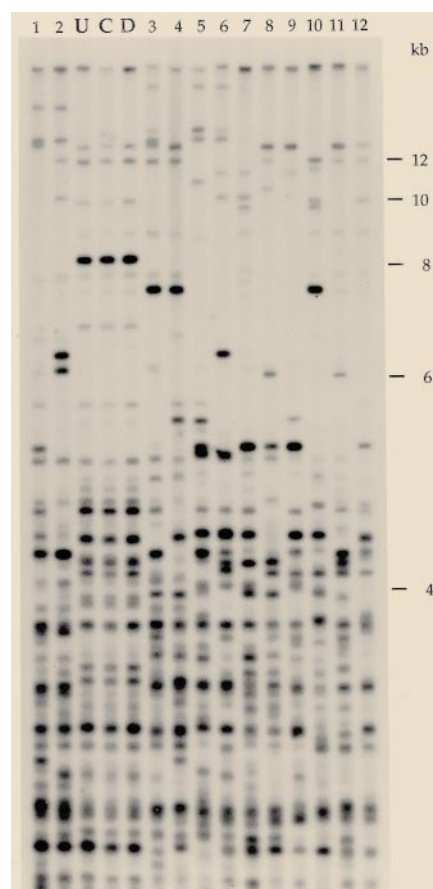


Figure 1 DNA fingerprint analysis of Dolly. DNAs were prepared from the donor udder (U), from the derived cell culture (C), and from blood from Dolly (D) and from control sheep 1–12. These DNAs were digested with *Mbo*I, electrophoresed through a 30-cm long 0.8% agarose gel, and Southern-blot hybridized⁶ with a cocktail of ³²P-labelled probes comprising porcine S0322 (ref. 14), human MS1 (ref. 11), M13 phage DNA⁷ and murine MMS10 (ref. 15). Samples 2 and 1 are from a ewe and her lamb, respectively. kb, kilobases.

data to calculate the probability that all 22 variable bands in the donor tissue would be present in an offspring of the donor¹⁰. This probability is low (8.6×10^{-5}) and is further reduced to about 3.5×10^{-7} by taking into account the lack of any paternal bands in this offspring that are not shared with the

donor tissue. Thus it is unlikely that Dolly was derived from a fetal cell.

To verify these findings, we retested all samples with a second cocktail of cloned human minisatellites MS40 (ref. 11), MS43 (ref. 11) and p λ g3 (ref. 12) plus a (GG A/T)_n repeat probe¹³. Highly variable profiles were obtained, with 14.8 ± 2.5 ($\pm$ s.d.) extra variable bands per sheep not detected by the first cocktail of probes. Again, Dolly, the mammary tissue and the cell culture were indistinguishable (data not shown). Inclusion of these additional polymorphisms in the statistical calculations above substantially lowered all probability values. We therefore conclude that Dolly is derived from the nucleus of a cell from the mammary gland of the adult donor.

Esther N. Signer, Yuri E. Dubrova,

Alec J. Jeffreys

Department of Genetics, University of Leicester,

Leicester LE1 7RH, UK

e-mail: est@le.ac.uk

Colin Wilde, Lynn M. B. Finch,

Michelle Wells, Malcolm Peaker

Hannah Research Institute, Ayr,

Scotland KA6 5HL, UK

e-mail: wildec@hri.sari.ac.uk

1. Wilmut, I. *et al.* *Nature* **385**, 810–813 (1997).
2. Sgarrella, V. & Zinder, N. D. *Science* **279**, 635–636 (1998).
3. Wilde, C. L., Addey, C. V. P., Boddy, L. M. & Peaker, M. *Biochem. J.* **305**, 51–58 (1995).
4. Finch, L. M. B. *et al.* *Biochem. Soc. Trans.* **24** (Suppl.), 369 (1996).
5. Hermans, I. F., Morris, C. A., Chambers, G. K., Towers, N. R. & Jordan, T. W. *Anim. Genet.* **24**, 385–388 (1993).
6. Signer, E. N. & Jeffreys, A. J. in *DNA Fingerprinting: State of the Science* (eds Pena, S. D. J., Chakraborty, R., Epplen, J. T. & Jeffreys, A. J.) 421–428 (Birkhauser, Basel, 1993).
7. Vassart, G. *Science* **235**, 683–685 (1987).
8. Pena, S. D. J., Chakraborty, R., Epplen, J. T. & Jeffreys, A. J. (eds) *DNA Fingerprinting: State of the Science* 293–459 (Birkhauser, Basel, 1993).
9. Jeffreys, A. J. & Morton, D. B. *Anim. Genet.* **18**, 1–15 (1987).
10. Jeffreys, A. J., Turner, M. & Debenham, P. *Am. J. Hum. Genet.* **48**, 824–840 (1991).
11. Wong, Z., Wilson, V., Patel, I., Povey, S. & Jeffreys, A. J. *Ann. Hum. Genet.* **51**, 269–288 (1987).
12. Wong, Z., Wilson, V., Jeffreys, A. J. & Thein, S. L. *Nucleic Acids Res.* **14**, 4605–4616 (1986).
13. Armour, J. A. L., Croisier, M., Malcolm, S., Chan, J. C.-T. & Jeffreys, A. J. *Proc. R. Soc. Lond. B* **261**, 345–349 (1995).
14. Signer, E. N., Gu, F. & Jeffreys, A. J. *Mamm. Genome* **7**, 433–437 (1996).
15. Bois, P., Williamson, J., Brown, J., Dubrova, Y. E. & Jeffreys, A. J. *Genomics* **49**, 122–128 (1998).

A cellulase gene of termite origin

The traditional view of cellulose digestion in animals is that they cannot produce their own cellulase, and so rely on gut microorganisms to hydrolyse cellulose. A classic example of this symbiosis is that between phylogenetically lower termites and the unicellular organisms (protists) that colonize their hindguts: cellulose fermented to acetate by the protists can be used as an

energy source by the termite¹. There is evidence for the production of endogenous cellulase components by termites and other wood-feeding insects²; however, an unambiguous origin for such enzymes¹ has not been established, to our knowledge, until now. Here we describe the first insect cellulase-encoding gene to be identified, *RsEG*, which encodes an endo- β -1,4-glucanase (EC 3.2.1.4) in the termite *Reticulitermes speratus*.

Using antiserum raised against an endo- β -1,4-glucanase purified from *R. speratus*³, we screened a recombinant phage comple-

mentary DNA library from this species and identified a partial sequence encoding a peptide with similarity to cellulases from glycosyl hydrolase family 9 (GHF9) (ref. 4). We then obtained the complete coding region of *RsEG* by rapid amplification of complementary DNA ends.

Although the source of messenger RNA for this study was the salivary glands, a part of termites that lacks microorganisms², we completely confirmed the endogenous origin of the gene by Southern blot analysis of DNA extracted from degutted termites (results not shown) and with amplification

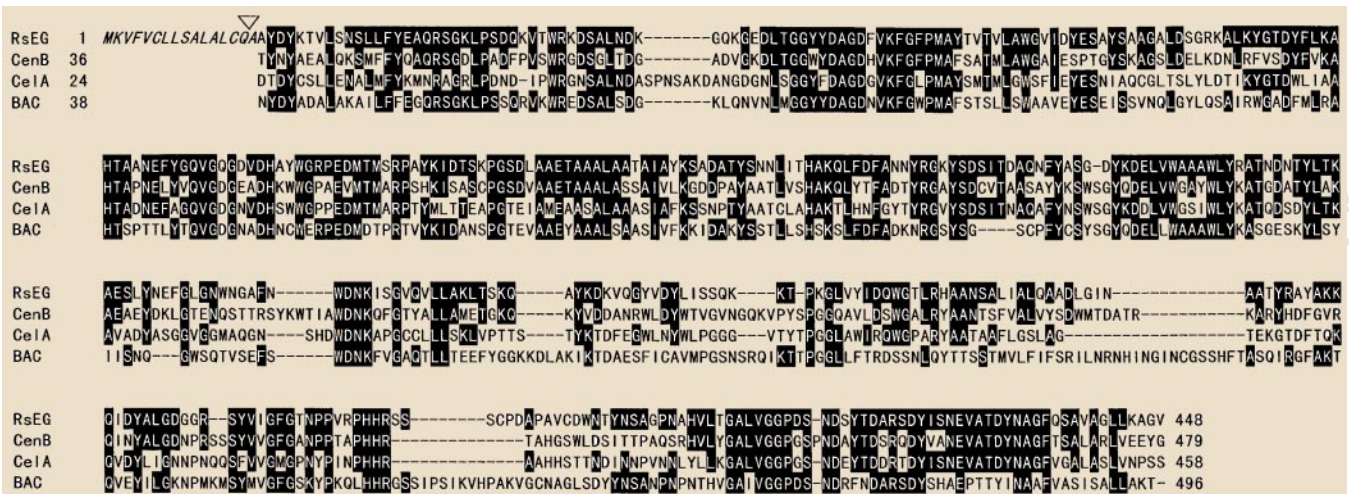


Figure 1 Predicted RsEG amino-acid sequence aligned with bacterial (*Cellulomonas fimi* CenB), protist (*Dictyostelium discoideum* CelA) and plant (*Phaseolus vulgaris* bean abscission cellulase (BAC)) members of GHF9. The alignment was done with the program Clustal W using mature RsEG. The putative 16-amino-acid RsEG signal peptide is shown in italics. The position of a 450-base-pair phase 1 intron present in an RsEG genomic fragment obtained by the polymerase chain reaction is indicated by an inverted triangle. Shaded amino acids show identity with RsEG. Coordinates have been deposited with the DNA Data Bank of Japan (RsEG, accession number AB008778), or with GenBank (CenB, CelA and bean abscission cellulase, accession numbers M64644, M33861 and U34754, respectively).

by the polymerase chain reaction of a partial RsEG genomic fragment. This fragment contains an exon sequence identical to that of RsEG cDNA and is interrupted by a 450-base-pair intron with typical eukaryotic splicing sites (Fig. 1).

The predicted mature RsEG protein shares 52%, 48% and 42% identity with the catalytic domains of selected bacterial, protist and plant members of GHF9, respectively (Fig. 1). The first 16 of the 448 amino acids of RsEG represent a secretory peptide, deduced from the amino-terminal sequences of two chromatographically distinct endo- β -1,4-glucanases purified from *R. speratus*³; these amino-terminal ends are identical. It is not clear at present which of these proteins is RsEG.

The amino-acid residues thought to be involved in catalysis and substrate-binding in all GHF9 members⁵ are conserved in RsEG, further confirmation that RsEG encodes an active cellulase.

A conspicuous feature of the termite endo- β -1,4-glucanase is the absence of any motif similar to the ancillary modules found in most bacterial and fungal cellulase components⁶. The presence of a cellulose-binding domain in such cellulases is thought to enhance activity towards crystalline cellulose⁷. RsEG appears to consist only of a single catalytic domain, a characteristic of plant endo- β -1,4-glucanases. However, unlike the latter enzymes, which have no apparent activity towards crystalline cellulose⁵, the termite endo- β -1,4-glucanase exhibits low exo- β -1,4-glucanase activity towards this substrate³, producing cellobiose and glucose. The specific activity of this reaction is similar in magnitude to that of bacterial endoglucanases from GHF9 (ref. 8).

Thus the endogenous cellulase system

of *R. speratus*, including an apparent β -glucosidase⁹, is secreted in the salivary glands and is capable of producing glucose from crystalline or amorphous cellulose present in ingested wood. Although this endogenous cellulase is not 'complete' (able to exhaustively digest native cellulose), the production of glucose from crystalline cellulose by this cellulase (estimated *in vitro*) can account for a substantial amount of the carbon dioxide expired by whole insects⁹.

Cellulose not hydrolysed in the anterior portion of the gut then travels to the hindgut, where it can be endocytosed by symbiotic protists and fermented to produce acetate, an important energy source and precursor for biosynthesis in the termite¹. Use of this dual cellulase system explains why wood-glucan assimilation in lower termites can be greater than 90% (ref. 1).

The evolutionary origin of cellulases in termites is intriguing. Although RsEG exhibits marked similarity to bacterial and protist members of GHF9, the fact that RsEG contains introns makes the prospect of recent horizontal gene transfer from these species unlikely. Indeed, the amino-terminal amino-acid sequences of two putatively endogenous cockroach endoglucanases¹⁰ both exhibit over 75% identity with the mature RsEG amino terminus. If these enzymes prove to be of insect origin, it is likely that an ancestral GHF9-like cellulase gene was present before the divergence of cockroaches and termites, an estimated 250 million years ago¹¹.

Hirofumi Watanabe, Hiroaki Noda

National Institute of Sericultural and Entomological Science, Tsukuba, 305-8634 Japan
e-mail: hinabe@affrc.go.jp

Gaku Tokuda

Bio-Oriented Technology Research Advancement

Institution, Omiya, 331-8537 Japan

Nathan Lo

Department of Biochemistry,
The University of Sydney, Sydney,
New South Wales 2006, Australia

1. Breznak, J. A. & Brune, A. *Annu. Rev. Entomol.* **39**, 453–487 (1994).
2. Slaytor, M. *Comp. Biochem. Physiol. B* **103**, 775–784 (1992).
3. Watanabe, H. *et al. Insect Biochem. Mol. Biol.* **27**, 305–313 (1997).
4. Henriissat, B. & Bairoch, A. *Biochem. J.* **293**, 781–788 (1993).
5. Brummell, D. A., Lashbrook, C. C. & Bennett, A. B. *Am. Chem. Soc. Symp. Ser.* **566**, 100–129 (1994).
6. Béguin, P. & Aubert, J. *FEMS Microbiol. Rev.* **13**, 25–58 (1994).
7. Tomme, P., Warren, R. A. J. & Gilkes, N. R. *Adv. Microbiol. Physiol.* **37**, 1–81 (1995).
8. Tomme, P., Kwan, E., Gilkes, N. R., Kilburn, D. G. & Warren, R. A. J. *J. Bacteriol.* **178**, 4216–4223 (1996).
9. Inoue, T., Murashima, K., Azuma, J. I., Sugimoto, A. & Slaytor, M. J. *Insect Physiol.* **43**, 235–242 (1997).
10. Scrivener, A. M., Zhao, L. & Slaytor, M. *Comp. Biochem. Physiol. B* **118**, 837–843 (1997).
11. Bandi, C. *et al. Proc. R. Soc. Lond. B* **259**, 293–299 (1995).

Arctic springtime depletion of mercury

The Arctic ecosystem is showing increasing evidence of contamination by persistent, toxic substances, including metals such as mercury¹, that accumulate in organisms. In January 1995, we began continuous surface-level measurements of total gaseous mercury in the air at Alert, Northwest Territories, Canada (82.5° N, 62.5° W). Here we show that, during the spring (April to early June) of 1995, there were frequent episodic depletions in mercury vapour concentrations, strongly resembling depletions of ozone in Arctic surface air, during the three-month period following polar sunrise (which occurs in March)^{2,3}.

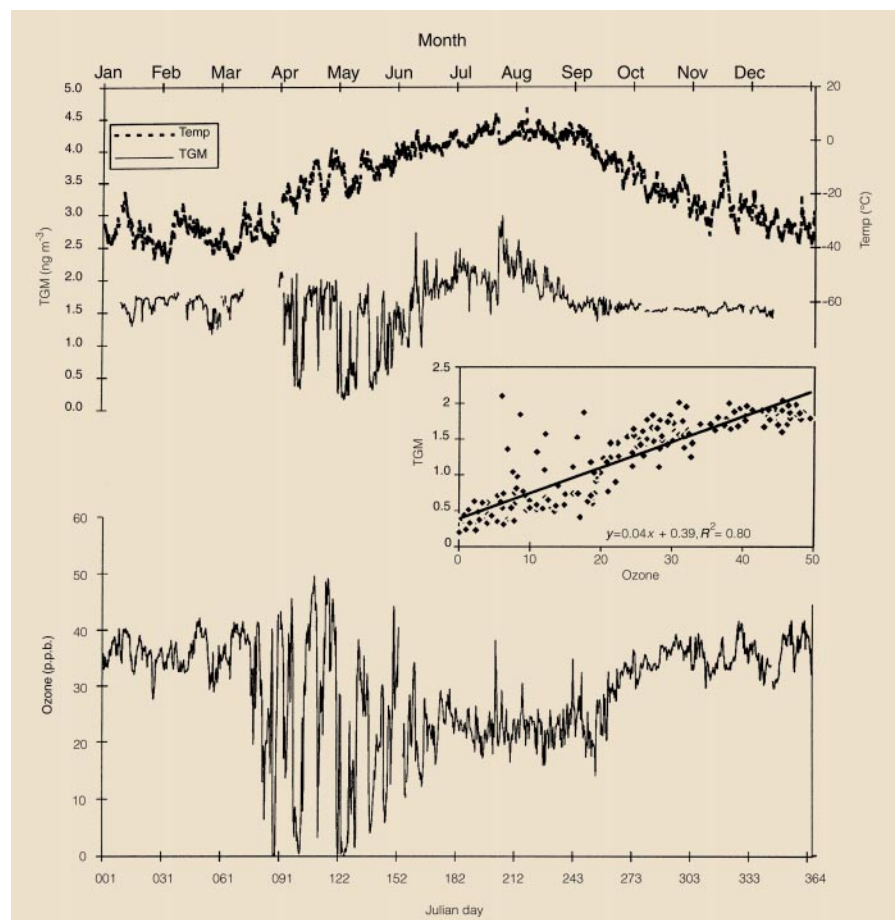


Figure 1 Time series of six-hour average values for air temperature and for total gaseous mercury (TGM) and ozone concentrations at Alert, Canada, in 1995. The inset shows concentrations of TGM versus ozone at Alert for the period from 9 April 1995 to 29 May 1995. (Note $R^2 = 0.8$ for the correlation between TGM and ozone concentrations during this period.)

Of the metals of environmental interest in the Arctic ecosystem, mercury is unique because it exists in the atmosphere predominantly in its gaseous state. This state, at least in remote locations, consists almost exclusively of elemental Hg^0 with a global atmospheric residence time of roughly 6–24 months. Thus, this species of mercury can undergo long-range atmospheric transport^{4,5} and the atmosphere represents a major pathway for its introduction into the Arctic environment.

Figure 1 shows the 1995 time series of six-hour average values for air temperature and concentrations of surface total gaseous mercury (TGM) and ozone at Alert. TGM levels show a distinctive pattern ('fingerprint') during each season. A dramatic increase in the variability of the mercury vapour and ozone signals is seen following polar sunrise, with a pronounced tendency towards unusually low concentrations. Before mid-March, TGM and ozone levels show only a modest ('normal') amount of variation and are found at the expected surface air concentrations, between 1 and 2 ng m^{-3} (~ 0.1 – 0.2 parts per trillion) and 30 to 50 parts per 10^9 (p.p.b.), respectively. However, during the three-month period

after polar sunrise, TGM levels frequently drop well below 1 ng m^{-3} , and those of ozone are often less than 10 p.p.b. and sometimes as low as the detection limit (0.5 p.p.b.) of the ozone analyser⁶.

TGM concentrations peak during summer. This may be due to temperature- and/or sunlight-induced emission or re-emission of volatile mercury species from the Earth's surface (terrestrial and/or aquatic ecosystems; on continental, regional, and/or local scales) during summer⁷. During the autumn, TGM concentrations at Alert are remarkably constant. Indeed, for a relatively long-lived air pollutant such as mercury, a small variation in its atmospheric concentration would be expected.

The springtime dry deposition of mercury can be estimated if the following assumptions are made. First, by analogy with Arctic tropospheric ozone depletion^{2,3,6,8}, the springtime depletion of mercury vapour occurs in the boundary layer of the atmosphere and the average height at which the boundary layer is 'capped' (because of temperature inversions) is $\sim 500 \pm 100$ m (ref. 9). Second, chemical conversion of Hg^0 yields either particulate-phase mercury or reactive gaseous mercury

species. Third, dry-deposition velocities, under environmental conditions associated with depletion events, are $\sim 0.1 \text{ cm s}^{-1}$ for particulate-phase mercury and $\sim 0.5 \text{ cm s}^{-1}$ for reactive gaseous mercury. Fourth, the air column within the boundary layer was 'emptied' of its mercury content at least five times in the 75-day period (from 1 April to 15 June 1995) during which mercury vapour was subject to depletion episodes.

We used the 'normal' mean springtime TGM concentration (the average TGM concentration during those times when mercury is not depleted) of 1.84 ng m^{-3} as a reference value, and, using the above assumptions, we have estimated (conservatively) with our data from Alert an average springtime dry-deposition flux for mercury of $2.5 \pm 0.5 \text{ ng m}^{-2} \text{ h}^{-1}$. This value is about twice the particle dry deposition flux of 1.2 $\text{ng m}^{-2} \text{ h}^{-1}$ estimated for north-central Wisconsin¹⁰, a relatively remote mid-continental setting for which the atmospheric transport and attendant mercury deposition are known to be important (from considerations of the mass balance of mercury). Note that the Wisconsin dry-deposition estimate was derived assuming a relatively high particle dry deposition velocity of 0.5 cm s^{-1} .

We conclude that springtime conversion of mercury vapour produces one or more mercury species with much shorter atmospheric residence time(s) than Hg^0 . This as yet undefined chemical oxidation mechanism provides an important environmental pathway for the introduction of mercury into the biosphere, thus potentially affecting large areas of the Northern Hemisphere at a time of the year when biota are preparing for peak summertime activity.

W. H. Schroeder, K. G. Anlauf, L. A. Barrie, J. Y. Lu, A. Steffen

Atmospheric Environment Service,
4905 Dufferin Street, Toronto,
Ontario M3H 5T4, Canada

D. R. Schneberger

Tekran Inc., 132 Rainside Road, Toronto,
Ontario M3A 1A3, Canada

T. Berg

Norwegian Institute for Air Research, PO Box 100,
N-2007 Kjeller, Norway

1. Canadian Arctic Contaminants Assessment Report (eds Jensen, J., Adare, K. & Shearer, R.) 460 (Indian and Northern Affairs Canada, Ottawa, Ontario, 1997).
2. Barrie, L. A., Bottenheim, J. W., Schnell, R. C., Crutzen, P. J. & Rasmussen, R. A. *Nature* **334**, 138–141 (1988).
3. Barrie, L. A. & Platt, U. *Tellus* **49B**, 450–454 (1997).
4. Schroeder, W. H. & Markes, J. J. *Great Lakes Res.* **20**, 240–259 (1994).
5. Schroeder, W. H. & Munthe, J. *Atmos. Environ.* **32**, 809–822 (1998).
6. Anlauf, K. G. *et al.* *J. Geophys. Res.* **99**, 25345–25353 (1994).
7. Xiao, Z. F., Munthe, J., Schroeder, W. H. & Lindqvist, O. *Tellus* **43B**, 267–279 (1991).
8. Pearce, F. *New Sci.* **154**, 17 (1997).
9. Bradley, R. S. & Keimig, F. T. *J. Geophys. Res.* **97**, 15699–15712 (1992).
10. Fitzgerald, W. F., Mason, R. P. & Vandal, G. M. *Wat. Air Soil Pollut.* **56**, 745–767 (1991).

The joy of secants

Trigonometric Delights

by Eli Maor

Princeton University Press: 1998. Pp. 236.

\$24.95, £17.95

Geometry Civilized: History, Culture, and Technique

by J. L. Heilbron

Oxford University Press: 1998. Pp. 309.

£35, \$65

Jeremy Gray

Mathematics in general, and geometry in particular, seem to their practitioners to be both essential and attractive aspects of life. Not so to many who struggled with them at school; and still less, one must presume, do today's school students value them, because they have almost lapsed from the syllabus. How to bridge this fissure?

These two books, which cover geometry and trigonometry, offer an interesting pair of attempts. As their titles suggest, John Heilbron's is the more substantial, and weaves in and out of a delight in the mathematics and a sensitive attention to the variety of cultural occasions for it. The book by Eli Maor is similar, but is lighter culturally and deeper mathematically. But both share a belief that doing the mathematics — solving problems — is part of the fun; both trace the origins of these ideas to a variety of cultures; and both glory in the achievements that geometry and trigonometry have made possible. Any reader not frightened of mathematics should be able to share that emotion and come away with a deeper appreciation of these subjects.

Heilbron's book has an unconventional organization that contributes greatly to its charm. Lofty themes, such as cathedral windows, and aesthetic gems, such as Islamic designs, are mixed with difficult (and sometimes perhaps contrived) exercises in a survey of real and imaginary fields. The exercises are culled from such sources as Egyptian papyri, *The Ladies' Diary* and textbooks down the ages; and Heilbron adds many of his own. They occur on almost every page, not hived off at the end where they can be skipped; reader, do them! Their effect is to humanize the matter, to connect the important with the mundane, and to suggest that only by mastering these challenges could geometers and architects have accomplished their greatest work.

Equally, the mixture of the elevated and the irritating makes the 'mere' problems more interesting, without implying that they are more significant in themselves. Modern, especially American, calculus textbooks like to insinuate that the subject is 'really' useful by similar devices, but they fail because the examples are all too often artificial questions dressed up to look like real ones. Heilbron



Polygon with polymaths: Euclid and his students construct a star hexagon, in this detail from Raphael's fresco *The School of Athens*.

succeeds because of the wealth of his examples. Surely, one comes to realize, this is how the old experts saw the subject: as a real mix of the sublime and the fiddly. Designing a large Gothic window and working out an inelegant trigonometric problem are part of the same activity, which a mathematician achieves by doing the sums.

The book is wonderfully illustrated. There are diagrams on almost every page, apt illustrations from older books, and well chosen photographs, together with eight colour plates. The appearance of the book is quite seductive, for which Oxford University Press should be congratulated.

Heilbron covers Euclidean geometry, including the behaviour of circles, and plane trigonometry. All is explained from scratch, but not in the style of a textbook. Nor is it exactly a history of these topics, although it is erudite. The effect is to impart dignity to the subject, and wisdom to its readers.

Maor's book is smaller and lighter, but none the worse for that. It includes topics drawn from the calculus of the trigonometric functions: infinite series, the hyperbolic trigonometric functions, and complex numbers. It is more expository than Heilbron's book, with fewer exercises, and is more conventionally organized as a history. Topics include Ptolemy's tables, the Hindu origin of the trigonometric functions, trigonometric

identities, the shape of the Earth, cartography, de Moivre's formula and its consequences, and rigour in analysis. Little of this is new, but the mixture is engagingly presented.

Students of trigonometry and their teachers will enjoy finding the core syllabus so refreshingly opened up. If they find the accounts, especially towards the end of the book, to be tantalizingly short, this is all the more reason to emulate the author and find out more for themselves, and Maor gives advice on how this can be done.

Both authors suggest that geometry and trigonometry were casualties of the 'new maths' reforms of the 1960s, and that the calculator-driven education of today is no better. The case for these subjects, at a genuinely elementary level, based on their utility and beauty, has seldom been better put. Heilbron's, in particular, is a deeply humanist work, which is a strange thing to say about a book with detailed geometry on every page. It may be that the devil is in the details, but so too is the step-by-step, line by painful line, progress of human achievement. The full immersion in the subject that Heilbron provides allows one to recognize oneself in the work of others, and to share in that inclusive web of our history, culture and technique. □
 Jeremy Gray is in the Faculty of Mathematics and Computing, The Open University, Milton Keynes MK7 6AA, UK.

Carnivores in living colour

The Big Cats and Their Fossil Relatives: An Illustrated Guide to Their Evolution and Natural History

by Mauricio Antón and Alan Turner
Columbia University Press: 1997. Pp. 256.
\$46, £31.95

Adrian M. Lister

In this attractive book, the artist Mauricio Antón has collaborated with the palaeontologist Alan Turner to produce a persuasive account of an extraordinary group of animals. Turner's text is informative and scholarly, yet entertaining and accessible. But Antón's illustrations are the revelation, and certainly rank among the best reconstructions of extinct animals ever produced.

Here are the big cats, engaged in all manner of activities: courting, hunting, feeding or simply watching. The animals are truly alive on the page, whether in pencil drawing or full-colour painting. Antón has observed living cats acutely, compared their detailed anatomy with those of their fossil relatives, and re-created with exceptional empathy the extinct forms. The ways in which the details of soft anatomy, such as ears and lips, can be deduced from skeletal anatomy, and the limitations of such reconstructions, are explained in the text and make interesting reading.

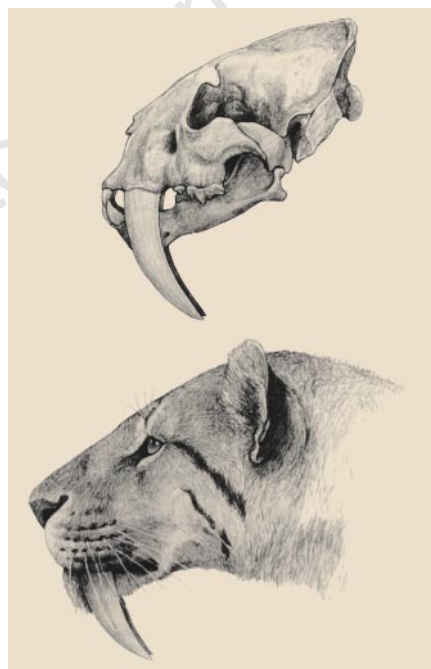
After a brief description of early forms that paralleled the true cats — nimravids, creodonts and marsupial 'cats' — the true cats are treated more fully. The text introduces us to living and extinct forms in the same format and with equal weight, a device that adds to the sense of the fossils as real animals and effectively conveys the full diversity of the group. Turner wisely avoids the common obsession of making phylogeny the central theme of all palaeontological work, concentrating instead on diversity and adaptation. He summarizes questions of relationship and nomenclature with sensible pragmatism, pointing out that "we still lack any clear idea of the immediate ancestors of the living groups or of the precise pattern of relationships between them".

The two main groups of big cats are the Felinae, which have conical canines, and the wholly extinct Machairodontinae, or sabre-tooths. The authors describe different adaptive types based on the known ecology and behaviour of modern forms, and use these to deduce the locomotion (climbing versus running) and hunting (stalking versus ambush) strategies of the extinct forms from their anatomy.

The function of the sabre-toothed cats' distinctive flattened canines, which could be as long as 28 centimetres, has been much



Double jeopardy: two young sabre-toothed cats risk their delicate upper canines in an ill-judged attack.



Fleshed-out feline: a reconstruction of the head of the American sabre-tooth *Smilodon fatalis*.

debated. Turner and Antón suggest that they were too delicate to be used on moving prey. Instead they argue that the sabre-tooths used their enormously powerful front quarters to pull the prey to the ground before biting. While the animal was held down to prevent it struggling, a carefully placed pierce or slash of the throat by the canines would have led to the prey bleeding to death, in contrast to the suffocating grip of modern big cats. The book's illustrations of the cats pursuing and dispatching their prey are especially wondrous, whether the cat is stalking with intent

gaze, turning in a cloud of dust or caught in mid-pounce.

Even aspects of social structure (solitary or communal) can be deduced for extinct species, based on hunting strategy (from skeletal proportions), brain size and the composition of the fossil assemblage. But the authors are realistic about the limitations of such conclusions, especially given the flexibility exhibited by modern species.

In the final part of the book, Turner and Antón try to draw conclusions about episodes of evolution and extinction through geological time, and about the environmental forces that may have driven them. For these more contentious issues, the popular treatment is perhaps less satisfactory. For the first time, one feels the need to see the raw data to be persuaded by the arguments. Supposed extinction 'events' at 5, 2.5 and 0.9 million years ago seem to be somewhat extended in time, from the information given. But some hypotheses, if treated as such, are intriguing. For example, it is suggested that, between 2 and 3 million years ago, global cooling triggered the spread of more open, savannah habitats in Africa. This led to selection for swifter-running herbivores, favouring the feline cats at the expense of the sabre-tooths, which went into decline.

This is an outstanding book for both palaeontologist and 'neontologist' to relish. It should convince even those sceptical of the value of palaeontology that we live within a thin skin of biodiversity underlain by an extraordinary richness, which has much to teach us about evolution and its products. □
Adrian M. Lister is in the Department of Biology, University College London, London WC1E 6BT, UK.

Instrumental insects

The Songs of the Grasshoppers and Crickets of Western Europe

by D. R. Ragge & W. J. Reynolds

Harley: 1998. Pp. 591. Book £65, CDs £25

Darryl T. Gwynne and Glenn K. Morris

Singing grasshoppers and crickets have inspired poets, and have even been adopted as caged pets in certain Asian and European countries. Now the calls of those insects found in western Europe can be identified using this nicely packaged acoustic field guide, which includes keys to species, colour plates and compact discs. The insects included are from three taxonomic families that share the abilities to sing and jump but are genetically quite distantly related: Acrididae comprises grasshoppers and locusts, Gryllidae includes the familiar crickets of field and hearth, and Tettigoniidae is composed of the bushcrickets (known as katydids in North America and Australia).

The songs featured are almost all produced by males. This is not investigator bias but a fact of nature: those familiar buzzes and chirps of night-time walks and hot summer afternoons are mating calls. The role of the calls and how call patterns are made and heard are covered in the book's introductory chapters. But it is a rather modest, selective review, which provides neither an effective explanation for the mechanical basis of sound patterns nor any speculation on the origins of their "bewildering variety" — a virtual Morse code for the 171 species. Why did unique song codes evolve? Was it to prevent mismatings between taxa, or was it simply a by-product of other selection pressures within species?

But the intention here is to describe song codes and use them to identify species, and in this the book succeeds. The 400-page database of song patterns (call spectra, which are of diagnostic value in their own right, are unfortunately omitted) is a useful source of characters that could be traced onto proposed evolutionary trees of the Orthoptera, allowing a reconstruction of how these signals may have evolved within lineages of species. The identification guide will certainly be handy for naturalists and ecologists. Equipped with this book, a decent set of ears and a tape recorder, researchers can now quickly sample the abundance of species in various habitats without the bother of collecting and identifying specimens. Such methods have already been used in studies of the effects of habitat fragmentation in western Europe on the metapopulation dynamics and diversity of bushcricket species.

The pace of insect identification will, however, be slowed by user-unfriendly

acoustical terms such as 'macrosyllable' and 'echeme'. Ragge and Reynolds defend their jargon as "indispensable", yet the book's title speaks of the "songs" of grasshoppers and crickets, not their echeme sequences. Other authors have described cricket calls using simple terms such as 'pulse', 'chirp', 'trill' and 'song'; and, let's face it, 'diplosyllables' has none of the charm of the 'chirps' of Charles Dickens' *The Cricket on the Hearth*.

Despite the unhelpful terminology, these singing insects can be identified using the acoustic keys and tables in this book. (We tested them on a couple of European songsters that have invaded Canada.) Books that use sounds to identify regional singing faunas (which also include frogs, birds and cicadas) deserve a place on the naturalist's bookshelf alongside the more typical pictorial field guides. When packages such as this one, comprising books and compact discs, are available for more regions of the world, perhaps the identification of animals using sounds might one day rival the more traditional pastimes of spotting birds and butterflies using field marks. □

Darryl T. Gwynne and Glenn K. Morris are in the Zoology Department, University of Toronto in Mississauga, Mississauga, Ontario L5L 1C6, Canada.

Holy prehistory

From Black Land To Fifth Sun: The Science of Sacred Sites

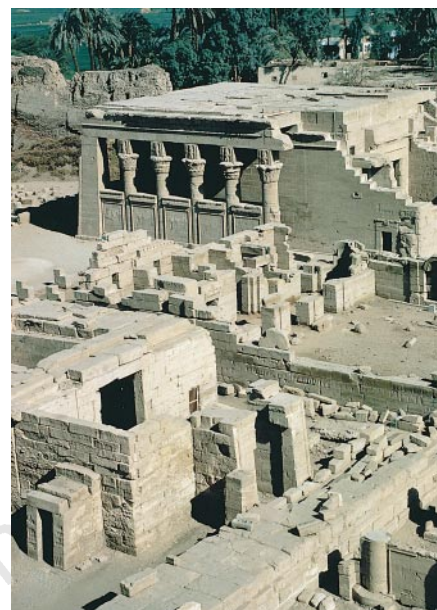
by Brian Fagan

Addison-Wesley: 1998. Pp. 403. \$26

Nicholas J. Saunders

The more science we throw at deciphering the past, the more apparent it becomes how different our ancestors were from us. Fittingly for an end-of-millennium perspective, scientific advances have combined with theoretical debates and philosophical sensitivities to enable archaeologists to appreciate the many different ways of 'being human in the world'. The challenge now is to acknowledge the power of technology to provide our data, yet avoid in our interpretations the modern rationalism that underpins that technology. This balancing act lies at the heart of today's multidisciplinary study of the past.

One of the most important archaeological advances of the past two decades has been the identification and understanding of the archaeological correlates of ideological activity, particularly in relation to ancient cosmologies, sacred sites and symbolic landscapes. Previously dominated by fantasy, fringe and 'new age' theories, this area has seen archaeology borrow ideas and insights from anthropology to enrich its range of interpretive models. Such has been the pace with which archaeology has embraced the hitherto poisoned chalice of investigating



Ritual site: temple at Dandarah, Egypt, dedicated to Hathor, the goddess of the sky and fertility.

'ritual behaviour' that it was only a matter of time before a popular synthesis appeared. In *From Black Land to Fifth Sun*, Brian Fagan marshals a selection of the evidence and presents his views on the science of sacred places.

Fagan's status as the dean of popular archaeology writers, his gift for making sense of jargon, and the nature of the topic promise a quality book, and that is almost what we get. One might have expected a seamless account of the many startling and diverse discoveries and insights of archaeology; instead it is an uneven set of case studies chosen and linked by virtue of their 'obviously' religious nature.

Part of the problem lies in the opening chapter, which oscillates between, for example, the hard science of mitochondrial DNA and the softer anthropological evidence for shamans, transformations, hallucinogenic visions and sacred space. Bearing in mind the range of examples that follow, a much stronger and more coherent statement was needed here.

Weaving together such disparate topics as Palaeolithic cave art, Neolithic Turkey, Stonehenge, Egypt, Minoan Crete and the Classic Mayan and Aztec civilizations of Mesoamerica was never going to be easy. Even



Bison on the rocks: a Palaeolithic cave painting from Altamira in Spain.

experts disagree about how the evidence should be interpreted. Fagan's attempts to use scientific methods to unlock non-western world views are not always successful. Although the linked chapters on current thinking about Palaeolithic cave art and ethnographic San (Bushman) art from South Africa work well, it is all too easy (and misleading) to think of the former as simply a more ancient version of the latter. Transforming shamans, magical visions and supernatural animals may be more attractive than old ideas such as 'sympathetic hunting magic', but they remain assumptions for all that.

Nevertheless, the first part of the book is effective: the stories are fresh, the insights dramatic, and Fagan tells the unfolding drama with verve. The chapter about the power of ancestors among the Dande of central Africa sparkles, mainly because of the author's direct involvement with archaeological investigations there. This is Fagan at his best.

But the narrative loses tension and focus as soon as it touches Stonehenge. Perhaps this most famous of megalithic monuments is just too familiar and over-published for a brief chapter to grab attention. Whatever the reason, the book's momentum and novelty fall away sharply from this point. Despite an interesting vignette of Colin Renfrew's work on the archaeology of cult at the Aegean site of Phylakopi, and an erudite and well-rounded account of North American Moundbuilders, there is little in the second half of the book to compare with the first. A lack of sharpness, some throwaway 'I was there' phrases and a sense of skimming across important issues do not help.

Towards the end, one could almost sense the publisher's deadline approaching. The chapter on the Maya is far too cursory, and the penultimate offering on Aztec Mexico contains nothing that has not been published many times before in popular books. An account of the 30-year-old project to map the pre-Aztec metropolis of Teotihuacan, a standard account of Aztec history, warfare and sacrifice, and a few pages about the excavations of the late 1970s at the Great Aztec Temple, peppered with snippets on sacred landscapes, do not really gel. The material is startling indeed, yet for some reason the account fails to connect.

Fagan has spotted a gap in the market, and with considerable skill has positioned this book accordingly. Its topicality, wide-ranging examples and lack of serious competitors make it a welcome introduction to what he calls "the archaeology of the intangible". Nevertheless, I was left with a sense that, given a little more time, this could have been so much better. □

Nicholas J. Saunders is in the Department of Anthropology, University College London, Gower Street, London WC1E 6BT, UK.

Is anybody out there?

Other Worlds: The Search for Life in the Universe

by Michael D. Lemonick
Simon & Schuster: 1998. Pp. 272. \$25

Alien Life: The Search for Extraterrestrials and Beyond

by Barry Parker
Plenum: 1998. Pp. 254. \$27.95, £16.94

Astronomical and Biochemical Origins and the Search for Life in the Universe: Proceedings of the 5th International Conference on Bioastronomy (IAU Colloquium 161)

edited by C. B. Cosmovici, S. Bowyer and D. Werthimer
Editrice Compositori: 1997. Pp. 814. \$100, Lire 160,000 (pbk)

David W. Hughes

The progress of science depends not only on asking the right question, but also on having the technical ability to make a reasonable attempt at producing an answer. Here the question is: "Are we alone in the Universe?". There are three possible answers: "Yes", "No" and "Don't know". At the end of the twentieth century we are still firmly stuck with the last one. Moving to either of the other possibilities would be one of the most significant steps in the history of scientific endeavour.

The answers "Yes" and "No" would be equally amazing, but the former would never be accepted. The Universe is so vast, and has already existed for so long, that we would never prove to everyone's satisfaction that a scintilla of life does not exist beneath some far-off stone, or on a planet behind one of the billions of distant stars.

But if we are not alone in the Universe, why have our co-inhabitants not been to see us, or at least phoned? Why is the Earth not littered with the remnants of crashed interstellar probes? Why are our skies not alive with the cross-talk of millions of interstellar communications? Our technological sophistication is not a serious problem here; the existence of extraterrestrial life should be obvious just by looking and listening.

If it is assumed that life is ubiquitous, we then have to ask if it is all based around the chemistry of carbon, and whether water is the universal solvent. Is the creation of life inevitable wherever it is warm and wet for aeons? Does all universal life depend on DNA, RNA and protein? And does it always take about 2.5 billion years for blue-green algae to evolve into *Nature* readers?

Why are there millions of living species on Earth but only one species of high intelligence? Does some extraterrestrial life also become intelligent, inquisitive and techno-

logical? Does it also develop the potential for self-destruction? And, assuming that life does exist out there, where does life on Earth rank in the intelligence and development league table of the Universe?

These and other bioastronomical questions have taken on greater urgency with the recent discovery of a host of planets around nearby Sun-like stars, hints that life could have existed on our neighbouring planet Mars, and the realization that conditions on Jupiter's moon Europa might be conducive to microbial life. This urgency is spawning conferences and books.

Astronomical and Biochemical Origins and the Search for Life in the Universe is a collection of 84 papers and reviews that were given in July 1996 at a bioastronomy conference in Capri, Italy. The topics include organic material in interstellar clouds, comets and meteorites; the search for planets around nearby stars and the form and distribution of these planetary systems; the origin and evolution of life and intelligence; and SETI, the search for extraterrestrial intelligence in the radio, infrared and optical wavebands. The author list is replete with the names of the famous, and to have three Nobel laureates at the same small meeting is most unusual. The papers are erudite but remarkably easy to understand: the interdisciplinary nature of this field forces researchers to keep jargon and obscurity to a minimum.

Other Worlds by Michael D. Lemonick and *Alien Life* by Barry Parker are both first-class, entertaining and authoritative examples of the art of scientific popularization; it is very difficult to choose between them. Both distil the contents of the bioastronomy enterprise and present a highly readable account of present progress and possible future developments.

Lemonick spent a great deal of time travelling to observatories and laboratories and talking to the astronomers and other scientists who are working at the frontiers of this fascinating subject. His account is full of anecdotal insights.

Parker, on the other hand, concentrates on the science and ideology. He contrasts panspermia with spontaneous generation, reviews the various forms of the anthropic principle, considers the influence of the discovery of extremophiles on the extent of stellar habitable zones, re-examines the Drake equation, and balances the expected waves of galactic colonization against the perceived absence of the 'calling cards' of extraterrestrial visitors. His speculation is firmly based on sound, well explained scientific principles, and he encourages us to prepare for the inevitable — and not too distant — conversation with an alien. □

David W. Hughes is in the Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, UK.

The structural basis of the activation of Ras by Sos

P. Ann Boriack-Sjodin*, S. Mariana Margarit†, Dafna Bar-Sagi† & John Kuriyan*‡

* Laboratories of Molecular Biophysics, and ‡ Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA

† Department of Molecular Genetics and Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

The crystal structure of human H-Ras complexed with the Ras guanine-nucleotide-exchange-factor region of the Son of sevenless (Sos) protein has been determined at 2.8 Å resolution. The normally tight interaction of nucleotides with Ras is disrupted by Sos in two ways. First, the insertion into Ras of an α -helix from Sos results in the displacement of the Switch 1 region of Ras, opening up the nucleotide-binding site. Second, side chains presented by this helix and by a distorted conformation of the Switch 2 region of Ras alter the chemical environment of the binding site for the phosphate groups of the nucleotide and the associated magnesium ion, so that their binding is no longer favoured. Sos does not impede the binding sites for the base and the ribose of GTP or GDP, so the Ras–Sos complex adopts a structure that allows nucleotide release and rebinding.

The Ras proteins are highly conserved guanine-nucleotide-binding enzymes that couple cell-surface receptors to intracellular signalling pathways controlling cell proliferation and differentiation^{1,2}. Ras acts as a molecular switch by cycling between active GTP-bound and inactive GDP-bound states. Whether GDP or GTP is bound to Ras is determined by the action of two classes of regulatory proteins: guanine-nucleotide-exchange factors, and GTPase-activating proteins. Exchange factors promote the activation of Ras by catalysing the exchange of GDP for GTP, whereas GTPase-activating proteins control the conversion of Ras to the inactive state by stimulating the hydrolysis of GTP to GDP².

Cell-surface receptors that signal through tyrosine kinases activate Ras by stimulating the guanine-nucleotide exchange reaction^{3–5}. Genetic and biochemical studies have indicated that this

reaction is controlled by the Ras guanine-nucleotide exchange factor Son of sevenless (Sos)⁶. Following ligand binding, Sos is brought from the cytoplasm to the activated receptor in a phosphotyrosine-dependent manner through adapter proteins such as Grb2. Grb2 contains SH3 domains that are bound constitutively to a carboxy-terminal proline-rich region of Sos, and the Grb2–Sos complex is recruited to activated receptors by interactions between the SH2 domain of Grb2 and phosphotyrosine residues on the receptor⁷. Because Ras is localized to the membrane, receptor activation results in an increase in the effective concentration of Sos in the vicinity of Ras, facilitating the exchange of bound guanine nucleotide for free cellular guanine nucleotides.

The region of Sos (relative molecular mass of about 150,000; $M_r \sim 150K$) that is functional for nucleotide exchange on Ras spans

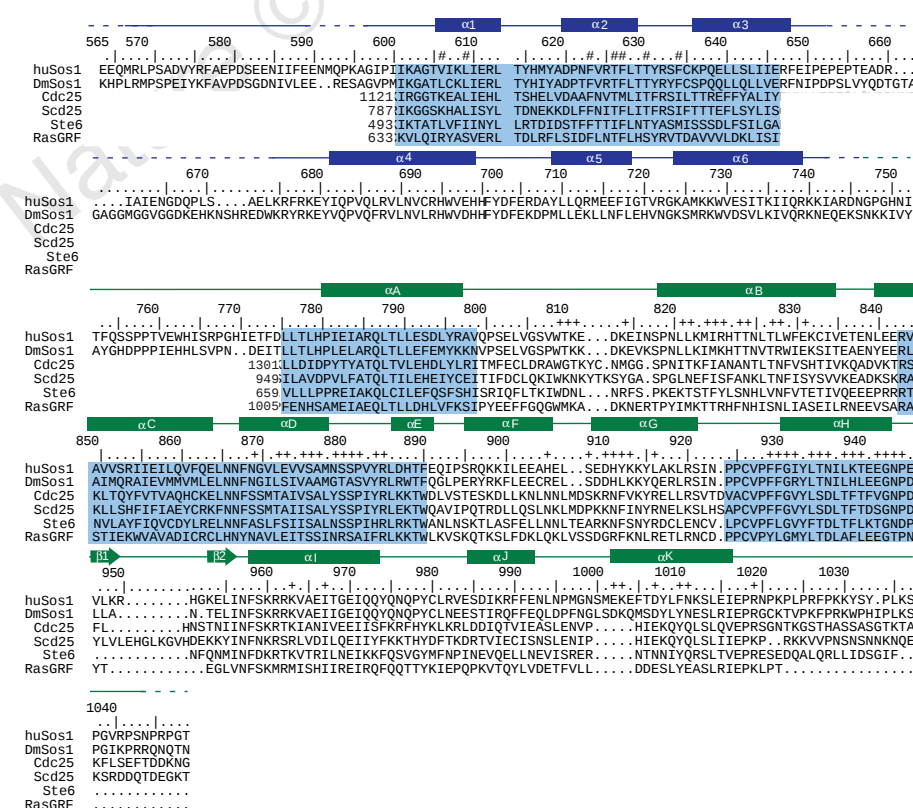


Figure 1 Sequence alignment of Ras-binding exchange factors with residue numbers of human (hu) Sos1 indicated. The secondary structure elements (α -helices are shown as rectangles, β -sheets as arrows, coil regions as a solid line, disordered residues as broken lines) of the N-domain are dark blue, and those of the catalytic domain are green. Conserved regions are shaded with pale blue. Residues of Sos at the Ras interface are indicated with +; residues in the N-domain that form the hydrophobic core with the catalytic domain are indicated with #. Sequences have been omitted where the sequence similarity between exchange factors is low. DmSos1, *Drosophila melanogaster* Sos1.

Table 1 Data collection, structure determination and refinement statistics

Data set	Resolution (Å)	Observations (total/unique)	R_{sym}^* (%)	Completeness (%)	$R_{\text{iso}}^\dagger$ (%)	No. of sites	Phasing power‡
Native	2.8	283,114/26,562	5.3 (28.7)	99.3 (99.9)			
PCMB§	3.1	616,299/19,766	7.7 (28.9)	93.4 (87.9)	22.8	5	1.68
Baker's dimercual§	3.25	640,979/16,988	7.6 (19.3)	96.9 (98.6)	21.5	6	0.80
EMTS§	3.25	634,455/17,174	7.3 (35.1)	95.4 (97.2)	20.0	4	1.94
Trimethyl lead acetate§	2.9	523,356/23,842	5.9 (27.0)	98.2 (99.3)	22.5	6	0.77
Selenomethionine	3.0	543,633/21,735	6.5 (29.7)	96.1 (93.2)	15.4	10	1.44
Gold cyanide§	3.05	368,880/20,927	5.4 (28.6)	98.3 (96.5)	33.5	3	1.46
Platinum terpyridine	3.25	240,941/16,915	10.6 (36.7)	98.9 (99.3)	18.5	6	1.35
Osmium chloride	3.1	202,555/20,025	6.8 (29.4)	99.6 (99.9)	23.5	4	1.55
PCMB/trimethyl lead acetate§	3.3	571,428/16,485	7.8 (30.0)	99.8 (99.5)	26.6	9	1.91

Overall figure of merit§ = 0.70

Data set	Resolution (Å)	Reflections	Total atoms	R -factor/ $R_{\text{free}}^\ddagger$	r.m.s.d. from ideal values	
					bonds (Å)	angles (°)
Native	30–2.8	26,502	5,010	0.222/0.281	0.007	1.26

* $R_{\text{sym}} = 100 \times \Sigma |I - \langle I \rangle| / \Sigma I$, where I is the integrated intensity of a given reflection. For R_{sym} and completeness, numbers in parentheses refer to data in the highest resolution shell.

† $R_{\text{iso}} = 100 \times \Sigma |F_{\text{PH}} - F_{\text{P}}| / \Sigma F_{\text{P}}$, where F_{PH} and F_{P} are the derivative and native structure factor amplitudes, respectively.

‡ Phasing power = $\Sigma |F_{\text{PH}} - F_{\text{P}}| / \Sigma |F_{\text{PH}} - F_{\text{P}}|$.

§ Anomalous data were used in the phasing of these derivatives.

¶ Figure of merit = $\langle \Sigma P(\alpha) e^{i\alpha} / \Sigma P(\alpha) \rangle$, where α is the phase and $P(\alpha)$ is the phase probability distribution.

‡ R -factor = $\Sigma |F_{\text{O}} - F_{\text{C}}| / \Sigma F_{\text{O}}$; R_{free} was calculated with 5% of the data.

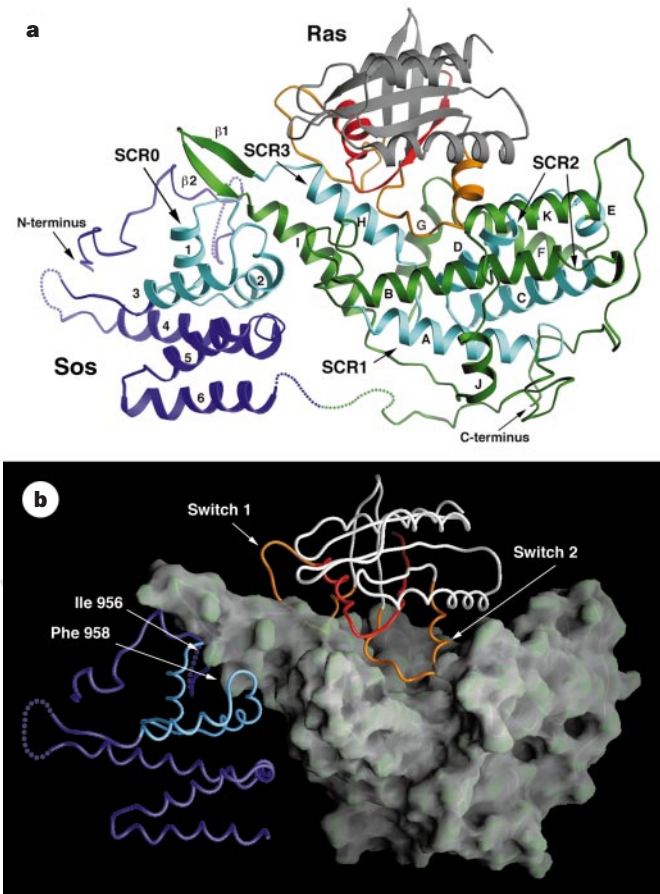


Figure 2 The complex of human H-Ras with the exchange-factor region of human Sos1. **a**, The N-domain of Sos (residues 568–741) is shown blue; the catalytic domain (residues 752–1044) is green; the Switch 1 and Switch 2 segments and the P-loop region of Ras (as defined here) are orange and red, respectively; conserved regions (SCRs) among Ras-family exchange factors are cyan²⁷. Disordered residues of Sos are shown as dotted lines. This and all other ribbon diagrams were generated using RIBBONS⁴⁴. **b**, The Ras-Sos complex is shown with the catalytic domain of Sos depicted as a molecular surface. Conserved residues Ile956 and Phe958 in the catalytic domain that form a hydrophobic interface with the N-domain are labelled. This and all other figures with molecular surfaces were generated using GRASP⁴⁵.

about 500 residues, and contains blocks of sequence that are conserved in other Ras-specific nucleotide exchange factors such as Cdc25, Sdc25 and Ras guanine-nucleotide-release factor (GRF)² (Fig. 1). However, in contrast to the high degree of structural conservation seen in the GTPases of the Ras family and others¹, there are distinct families of nucleotide exchange factors for the various guanine-nucleotide-binding protein (G protein) families that are structurally unrelated to each other^{8–13}. The structure of only one nucleotide exchange factor bound to its cognate G protein has yet been determined: that of ribosomal EF-Tu bound to its exchange factor EF-Ts^{12,13}.

To clarify the molecular mechanism of the activation of Ras by Sos, we have determined the crystal structure of the complex of a C-terminally truncated form of human Harvey-Ras (residues 1–166, hereafter referred to as Ras) with the guanine-nucleotide-exchange-factor region of human Sos1 (residues 564–1049, hereafter referred to as Sos). The structure reveals that Sos interacts extensively with Ras and stabilizes it in a nucleotide-free state by displacing the residues that coordinate the magnesium ion and the phosphate groups of the nucleotide and by partly occluding the magnesium-binding site. The structure also suggests a pathway for the rebinding of nucleotides to Ras, with consequent release of Sos, a process that is crucial for Sos to function as a nucleotide exchanger rather than a binding inhibitor.

General features of the structure

Ras and Sos form a tight complex in the absence of nucleotides. Crystals of this complex were obtained with one Ras–Sos complex (M_r 75K) in the asymmetric unit. The structure was determined by multiple isomorphous replacement and the molecular model was refined against data to 2.8 Å, resulting in a crystallographic R -value of 22.2% (free R -value of 28.1%; Table 1). The final model includes 439 residues of Sos (spanning residues 568–1044), 166 residues of Ras and 26 water molecules. There is no nucleotide or magnesium present in the crystals. The structure of Ras resembles that of the nucleotide-bound forms described previously^{19,46}, except for localized changes in the vicinity of the nucleotide-binding site. The structure of Sos, described below, seems to be unrelated to that of any other protein of known structure, based on a DALI¹⁴ search of the protein data bank.

The structure of the Ras exchange-factor region of Sos consists of two distinct α -helical structural domains (Fig. 2). The amino-terminal domain (N-domain; residues 568–741, α -helices α 1– α 6) does not interact with Ras, and seems to have a purely structural

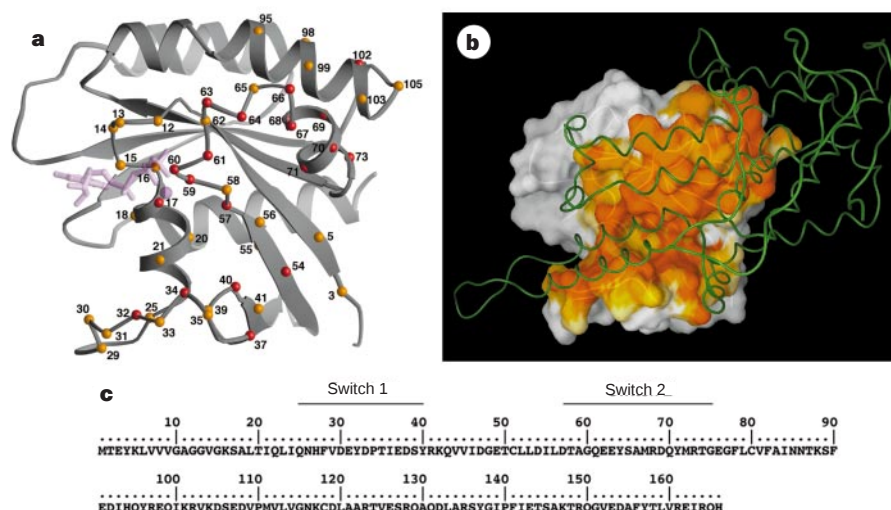


Figure 3 Interface surfaces of the Ras–Sos complex. **a**, Residues of Ras that form direct interactions with Sos are shown as red spheres; additional residues at the interface are orange spheres. The nucleotide is shown for reference only and is not present in the structure. **b**, The interface surface of Ras; the orientation is the

same as in **a**. The surface is coloured using a gradient: bright orange indicates atoms <4 Å from Sos, white indicates atoms >7 Å from Sos, lighter shades of orange indicate intermediate distances. Sos (N-domain deleted) is shown as a green ribbon. **c**, The primary sequence of Ras.

role. The C-terminal domain (residues 752–1044, α -helices α A– α K) contains within it all the residues that interact with Ras, and this region will be referred to as the catalytic domain. Analyses of the exchange-factor activity of Cdc25 and Sos have shown that the relatively well-conserved C-terminal catalytic domain suffices for catalytic activity^{15,16}. However, recombinant fragments of Sos that span the catalytic domain, but which lack some or all of the N-domain, are expressed poorly and have low solubility (data not shown), whereas the fragment used for the structure determination includes both domains and results in relatively high yields of soluble protein.

The structure of the catalytic domain consists of a series of helical hairpins that pack against each other (Fig. 2). A notable feature of the catalytic domain is the protrusion of a helical hairpin, formed by helices α H and α I, out of the core of the domain. Helix α H plays an important role in the nucleotide-exchange mechanism, and the main structural role for the N-domain appears to be the stabilization of the hairpin that presents this helix to Ras. Helices α 1 and α 2 of the N-domain together form a small hydrophobic groove into which two conserved hydrophobic side chains from helix α I of the catalytic domain are inserted (Ile 956 and Phe 958; Fig. 2b). The packing of these side chains into the α 1– α 2 groove, together with several adjacent interdomain interactions, is likely to be important for the stability and correct placement of the hairpin structure (Fig. 2).

The structure of the N- and catalytic domains of Sos is likely to be a good model for the general architecture of related guanine-nucleotide-exchange factors, such as Cdc25, Sdc25 and RasGRF (Fig. 1). Three regions of sequence conservation (structurally conserved regions, or SCRs) within the catalytic domain had been identified previously², and these are important either for the structural integrity of the domain (SCR1, helix α A and SCR2, helix α C) or for the interaction with Ras (SCR2, helix α D and SCR3; Fig. 2a). The region of the N-domain spanning helices α 1, α 2 and α 3 is highly conserved among Ras-specific nucleotide exchange factors (SCR0 in Fig. 2a)¹⁷. The hydrophobic nature of the groove between helices α 1 and α 2 is conserved, as are the residues on the catalytic domain that interact with the groove and the adjoining surface of the N-domain, suggesting that the interaction between the N-domain and the catalytic domain is conserved.

Structure of the Ras–Sos interface

The overall shape of the catalytic domain of Sos is that of an oblong bowl (Fig. 2), with Ras bound at the centre of the bowl. The regions of Ras that interact most closely with Sos include the phosphate-binding P-loop (residues 10–17) and surrounding segments (including strand β 1 and helix α 1), the Switch 1 region (defined here as residues 25–40) and the Switch 2 region (defined here as residues 57–75). Additional interactions are seen with helix α 3 (residues 95–105; Fig. 3a, b). The interface between Ras and Sos is primarily hydrophilic and very extensive, with 3,600 Å² of surface area buried in the complex. At the heart of the interface between Ras and Sos is a cluster of three hydrophobic side chains from the Switch 2 region of Ras (Tyr 64, Met 67 and Tyr 71) which are buried into the hydrophobic core of Sos at the base of the binding site. Surrounding this hydrophobic anchor is an array of polar and charged interactions between Sos and Ras that results in almost every external side chain of Switch 2 being coordinated by Sos (Fig. 4).

The most obvious effect of Sos binding to Ras is the opening of the nucleotide binding site as a result of the displacement of Switch 1 of Ras by the insertion of the helical hairpin formed by α H and α I of Sos (Fig. 5). An important consequence of the insertion of helix α H into the Ras active site is that it introduces a hydrophobic side chain (Leu 938), which blocks magnesium binding, and an acidic side chain (Glu 942), which overlaps the site where the α -phosphate of the nucleotide would otherwise be bound.

The Switch 2 region is held very tightly by Sos. Although this region tends to be poorly ordered in the nucleotide-bound forms of Ras^{19,46}, its conformation is very well defined in the Sos complex. The temperature factors for all atoms of Switch 2 are low (34 Å², compared with the average value of 49 Å² for all of Ras). The C-terminal end of Switch 2 is farthest from the nucleotide binding site, and side chains in this region interact with Sos but do not change their positions significantly. However, closer to the nucleotide-binding site the interactions between Sos and the side chains of Switch 2 result in a restructuring of the polypeptide backbone connecting β 3 and α 2 (Fig. 4a, b). This restructuring of the backbone is a crucial determinant of nucleotide exclusion, as it prevents the coordination of magnesium and phosphate (see below).

Switch 1 and Switch 2 are the only regions of Ras in which structural changes are directly induced by Sos. Comparison of the

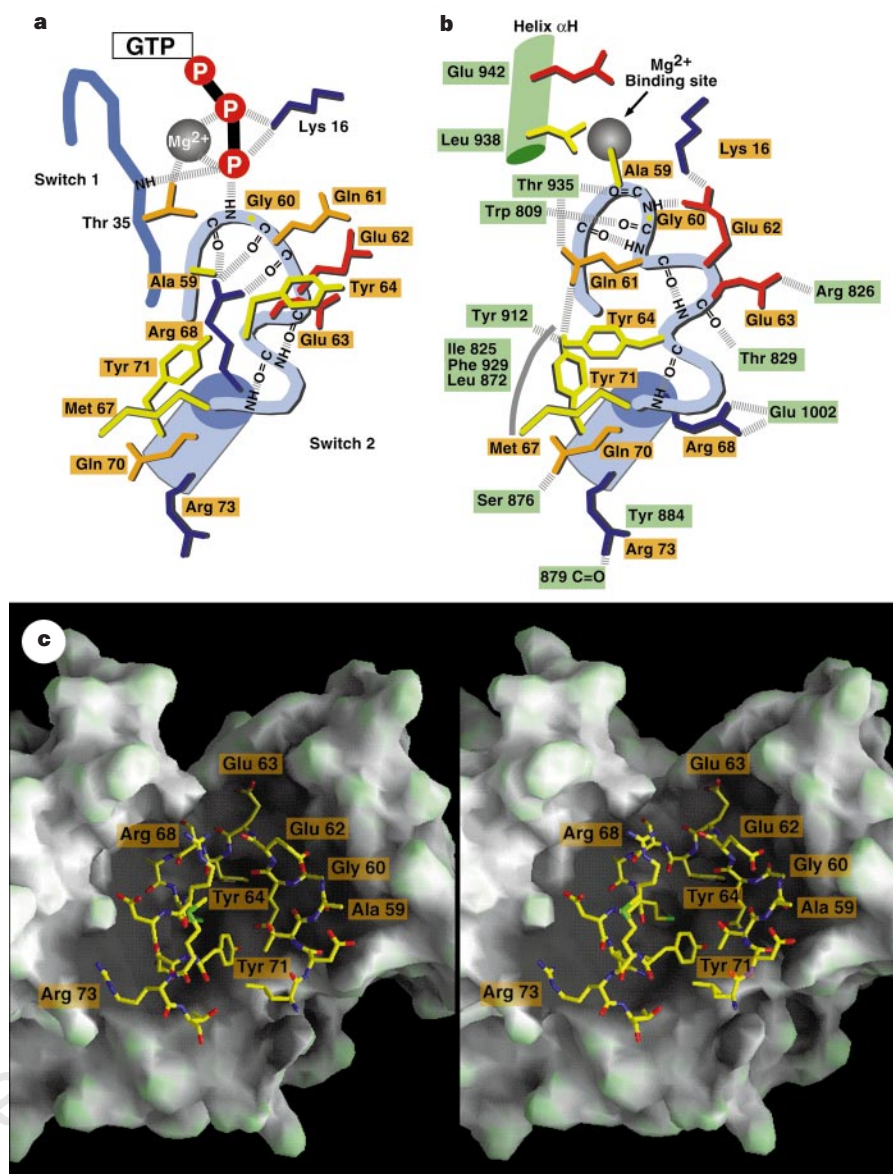


Figure 4 A schematic representation of the differences in the Switch 2 regions of a Ras–GTP analogue and Ras–Sos. **a**, Ras–GTP analogue (5P21)³⁹. **b**, Ras–Sos. Selected polar interactions are shown as broken lines, hydrophobic interactions

as solid arcs. **c**, A stereodiagram of the Switch 2 region of Ras superimposed on the surface of Sos. Selected residues of Ras are indicated.

structure of Ras in the Ras–Sos complex with that of the nucleotide-bound forms shows that there are also structural changes in one of the loops that binds the nucleotide base (loop 8, between $\beta 5$ and $\alpha 5$, residues 118–123) and the phosphate-binding P-loop. However, the changes in these loops seem to be a simple consequence of the absence of nucleotide.

Structural changes in the Switch 1 region

The change in the Switch 1 region of Ras when bound to Sos is drastic (Fig. 6). Switch 1 normally rises up from the end of $\alpha 1$ towards the P-loop region, to sandwich the nucleotide between it and the rest of Ras. Strand $\beta 2$ in Switch 1 forms an antiparallel interaction with strand $\beta 3$ in the central β -sheet of Ras. In the Sos complex, the C-terminal region of helix $\alpha 1$ is shortened by about one helical turn, the antiparallel β -sheet interaction with $\beta 3$ is completely disrupted, strand $\beta 2$ is partly melted, and Switch 1 is completely removed from the nucleotide-binding site.

One important aspect of the insertion of the helical hairpin of Sos into the Switch 1 region is that it does not result in a significant

occlusion of the guanine and ribose binding sites (Fig. 5d). Instead, this structural distortion breaks the network of direct and water-mediated interactions between Switch 1 and the nucleotide. For example, in the nucleotide-bound forms of Ras, Phe 28 interacts with the guanine base through a perpendicular aromatic–aromatic interaction (Fig. 5a). Mutation of Phe 28 to leucine results in a significant increase in the intrinsic rate of dissociation of nucleotide from Ras¹⁸. In the Sos complex, the C α of Phe 28 moves 9.6 Å and the side chain no longer interacts with the nucleotide-binding site (Fig. 5b).

Two side chains presented by helix αH of Sos, Leu 938 and Glu 942, directly impede the binding of magnesium and phosphate, respectively. The carboxylate group of Glu 942 of Sos is positioned near the location of the α -phosphate of GTP or GDP in the nucleotide-bound forms of Ras. In the Sos complex, Glu 942 forms a hydrogen bond with Ser 17 of Ras, a ligand of the magnesium ion in the nucleotide complexes of Ras, preventing the binding of the phosphates as well as the magnesium ion. Leu 938 further increases the hydrophobicity of the magnesium binding site

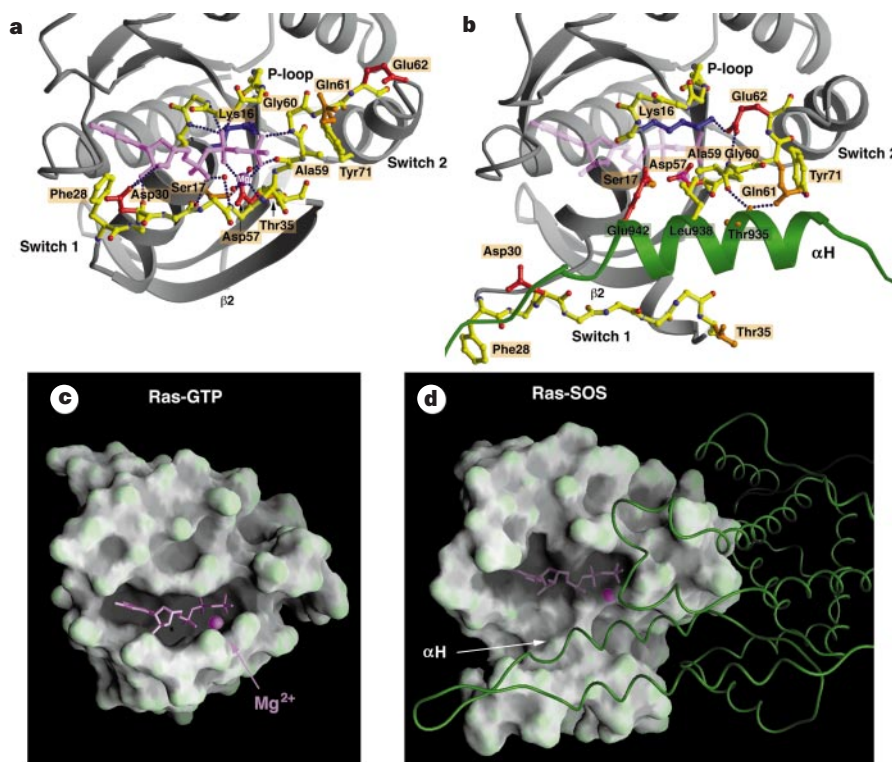


Figure 5 Interactions at the nucleotide-binding site. **a**, Selected interactions between Ras and a GTP analogue (5P21)³⁹. Water molecules are depicted as red spheres. **b**, Selected interactions between Ras and Sos are shown in the same orientation as in **a**. Only helix α H and selected side chains of Sos are shown. **c**, The nucleotide-binding site on the surface of Ras in a Ras-GTP analogue (5P21)³⁹

is shown. The side chain of Tyr 32 was deleted from the surface calculations for clarity. **d**, The surface of Ras in the Ras-Sos complex is shown with the backbone of Sos (N-domain deleted) as a green ribbon. Ras is in a slightly different orientation from that in **c**. In **b** and **d**, the nucleotide is shown for reference only.

in the Sos complex, which is also occupied by Ala 59 from Switch 2 (see below).

Changes in the Switch 2 region

The sensitivity of Switch 2 to the presence of either GTP or GDP is a consequence of the coordination of the terminal phosphate of GTP by the backbone amide nitrogen of Gly 60 (Figs 4a and 5a; refs 19, 46). In the following discussion, we compare the conformation of the Switch 2 region in the Ras-Sos complex to that seen in the GTP-bound form of Ras, because the Switch 2 region is not well ordered in the GDP-bound form¹⁹. The Switch 2 region of Ras makes important interactions with GTP and not with GDP^{19,46}. Nevertheless, structural changes that are induced in Switch 2 by Sos result in the exclusion of both GDP and GTP, because they affect magnesium binding as well as the conformation of Lys 16 in the P-loop, a crucial phosphate ligand.

Three specific features of the Switch 2 conformation are correlated with the disruption of nucleotide binding. First, the backbone conformation in the central region of Switch 2 is compressed in the Ras-Sos complex because of the formation of three consecutive β -turns (Fig. 4a, b). As a consequence of the first β -turn, the methyl side chain of Ala 59 is turned in towards the phosphate-binding site and occludes the position that would be occupied by the Mg^{2+} ion in nucleotide complexes. Second, the formation of the second β -turn results in the side chain of Glu 62 coordinating both the amide nitrogen of Gly 60 and the side chain of Lys 16, two residues involved in phosphate coordination (Figs 4 and 5). Finally, in the GTP-bound form, the polypeptide backbone of the β 3- α 2 loop adopts a conformation in which three carbonyl groups (that of 59, 60 and 61) are pointed inwards. The resulting anion hole coordinates the side chain of Arg 68, and positions the methyl group of

Ala 59 away from the magnesium-binding site. In the Sos complex, Arg 68 is removed from this internal location by interactions with Glu 1002 of Sos, and the anion hole is disrupted by the formation of the first two β -turns in the structure (Fig. 4a, b).

Comparison with EF-Tu and EF-Ts

The structure of only one other GTPase complexed to its exchange factor is yet known, that of EF-Tu bound to EF-Ts. EF-Tu is a GTPase that contains a nucleotide-binding domain that is topologically similar to Ras²⁰, as well as two additional domains. However, the mechanism of nucleotide release by its exchange factor EF-Ts is quite different from that seen here for Ras and Sos. The complex between EF-Tu and EF-Ts, like the Ras-Sos complex,

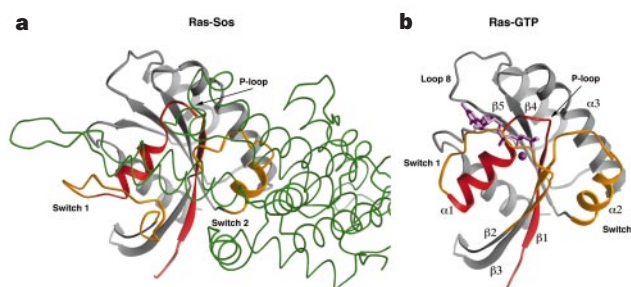


Figure 6 Comparison of Ras complexed with Sos and a GTP analogue. **a**, Sos; **b**, GTP analogue (PDB code 5P21)³⁹. The colouring scheme for Ras is the same as in Fig. 2. Sos (N-domain deleted) is shown as a green ribbon. Secondary structure elements of Ras important in nucleotide binding and Sos binding are labelled.

stimulates nucleotide release by disrupting the interactions of the phosphate groups of the nucleotide, leaving the binding site for the base and ribose unimpeded^{12,13}. Reorientation of a peptide bond in the phosphate-binding P-loop, induced by EF-Ts binding, results in the placement of a carbonyl oxygen in a position where it would collide with the β -phosphate of the nucleotide. In addition, global changes in the positions of the three domains result in conformational changes in the Switch 2 region that remove side chains that interact with the magnesium ion by means of water molecules. In contrast to Ras, the Switch 1 region of nucleotide-bound EF-Tu does not interact extensively with the nucleotide-binding site. Consequently, disruption of the Switch 1 structure does not seem to be a major component of the mechanism of EF-Ts action.

Implications for the reaction mechanism

Biochemical studies of Ras exchange factors have shown that the complex of Ras with these proteins is stable in the absence of nucleotides and is dissociated by the rebinding of either GDP or GTP^{16–18,21,22}. The principal role for the exchange factor is to facilitate nucleotide release, and it does not seem to control significantly the preferential rebinding of GTP over GDP^{16,22,23}. Cellular concentrations of GTP are ~ 10 -fold higher than GDP, which results in the loading of GTP onto Ras.

The mechanism of nucleotide release by the catalytic domain of murine Cdc25 (Cdc25^{Mm}) has been investigated recently using fluorescently labelled nucleotides¹⁶. The affinity of Cdc25^{Mm} for nucleotide-free Ras ($K_d = 4.6$ nM) is found to be several orders of magnitude higher than that for nucleotide-bound Ras, and the maximal acceleration by Cdc25^{Mm} of the rate of dissociation of nucleotide is more than 10^5 . Kinetic analysis of nucleotide association shows that the reaction proceeds by the formation of a ternary complex of a loosely bound nucleotide and Ras–Cdc25^{Mm} followed by conversion to a form in which the nucleotide is tightly bound to Ras¹⁶. In light of the structure of the Ras–Sos complex, the first step can be interpreted as the interaction of the base and the ribose of the nucleotide with the part of the Ras binding site that is not occluded by Sos. The second step would involve a conformational change in the Switch 2 segment and release of Switch 1, resulting in the restructuring of a competent binding site for phosphate and magnesium, and the subsequent dissociation of Sos. The mechanism by which Cdc25^{Mm} displaces the nucleotide does not depend solely on expulsion of the magnesium¹⁶. This is consistent with the Ras–Sos structure, as the mechanism involves interference with phosphate binding as well as displacement of the magnesium ion.

Analysis of Ras mutations

The structure of the Ras–Sos complex is consistent with several mutations in Ras that have highlighted the importance of the Switch 1 and Switch 2 regions in the interaction with nucleotide exchange factors^{18,24–30}. The importance of helix $\alpha 3$ (residues 102–105) has also been noted^{25,27,28}. The importance of Switch 2 for the recognition of the exchange factor is demonstrated by the analysis of mutations in residues that are not directly involved in nucleotide binding, but which affect GDP–GTP exchange. For example, mutation of Glu62 and Glu63 to histidine had no significant effect on the stability of the Ras–GDP complex¹⁸. However, both Ras mutants were severely compromised in their ability to be activated by Sdc25 (ref. 18). In the structure of the Ras–Sos complex, Glu62 and 63 of Ras are both seen to be crucial to the interaction with Sos (Figs 4b and 5b).

Of particular interest are dominant-negative mutants of Ras that appear to act by binding to and sequestering nucleotide-exchange factors^{31,32}. The most straightforward explanation of the action of these mutations is that they destabilize nucleotide binding^{22,32,33}, increasing the apparent affinity of Ras for Sos or other exchange factors. That the dominant-negative mutant Ras proteins act by binding to the exchange factor is also suggested by the fact that

mutations at residues important for the formation of the Ras–Sos interface result in a reversion of the dominant-negative behaviour^{26,29}. For example, in the yeast *Saccharomyces cerevisiae* protein Ras2, the mutation Ser 24 $\rightarrow$ Asn (corresponding to Ser 17 $\rightarrow$ Asn in human Ras) is dominant negative³². Substitution of Arg 80, Asn 81 (Arg 73, Thr 74 in human Ras) with Asp 80, Asp 81 in mutant (Ser 24 $\rightarrow$ Asn) Ras2 results in a loss of sensitivity to Sdc25 and reversion of the dominant-negative phenotype²⁹. In the Ras–Sos complex, Arg 73 (Arg 80 in Ras2) is involved in interactions with two residues of Sos (Fig. 4b), and mutation to Asp would clearly be disruptive.

Discussion

As a nucleotide-exchange factor, Sos functions under two apparently conflicting imperatives. The interaction between Sos and Ras must be strong enough to dislodge the tightly bound nucleotide, but the Ras–Sos complex must also be poised for subsequent displacement by incoming nucleotides. The structure of the Ras–Sos complex shows that Ras and Sos meet these demands by forming a tight complex that is anchored at one end of the nucleotide-binding site, where phosphate and magnesium are normally bound. The interface between Sos and Ras is mainly hydrophilic, suggesting a ready unzipping through water-mediated displacements of the coordinating side chains. The main interacting elements of Sos avoid direct occlusion of the nucleotide-binding site, except the region where the terminal phosphate groups and the magnesium ion are bound. This feature allows incoming nucleotides to reverse the process by competing for the groups that ligate the phosphate and metal ion.

The proliferative status of cells is critically dependent on the activation state of Ras. The structure of the Ras–Sos complex suggests at least two avenues for the design of inhibitors to block the activation of Ras by Sos. Nucleotide analogues that are designed to recognize the altered nucleotide-binding site in the Ras–Sos complex may help to stabilize the complex, mimicking the action of dominant-negative alleles of Ras. Alternatively, hydrophobic compounds that bind to the core hydrophobic region at the heart of the Ras binding site on Sos may effectively inhibit Sos action. $\square$

Methods

Expression and purification. *Escherichia coli* cells (BL21-DE3) were transformed with a pProEX HTb vector (Life Technologies) containing Ras (residues 1–166) linked to an N-terminal polyhistidine tag using the *Bam*HI and *Xho*I restriction sites. Protein production was induced with 250 μ M IPTG at a cell density of absorbance $A_{600} = 0.5$. Protein was expressed at 30 °C for 6 h. Cells were collected by centrifugation, resuspended in 20 mM Tris, pH 8.0, 300 mM NaCl at 4 °C, flash frozen and stored at –80 °C until needed. Once thawed, cells were lysed using a French press (EmulsiFlex-C5, Avestin), cell debris was removed by centrifugation, and the resulting cell lysate loaded onto a charged nickel binding column (HisBind; Novagen) pre-equilibrated with 20 mM Tris, pH 8.0, 500 mM NaCl and 20 mM imidazole. Protein was eluted using an imidazole gradient. Fractions containing Ras were pooled, dialysed into buffer A (20 mM Tris, pH 8.0, 100 mM NaCl), and concentrated. The polyhistidine tag was cleaved by tobacco etch virus (TEV) protease in buffer A and 2 mM β -mercaptoethanol at 4 °C for 48 h. After cleavage, protein was passed over a charged nickel-binding column pre-equilibrated with buffer A to remove uncleaved protein. Fractions containing pure Ras were pooled and concentrated. Expression and purification of Sos (residues 564–1049) was performed as above, with an additional purification step using a HiQ (Biorad) column pre-equilibrated with buffer A. Fractions containing Sos were concentrated in buffer A. The Ras–Sos complex was formed by incubating concentrated Sos with a 3- to 5-fold excess of Ras in buffer A for 1 h at 4 °C. Protein was loaded onto a Sephadex 75 gel filtration column (Pharmacia Biotech) pre-equilibrated with buffer A. Fractions containing complex were pooled and concentrated to 10 mg ml^{–1}. Approximately 30 mg purified complex could be obtained from 16 litres of *E. coli* cell culture.

Crystallographic analysis. Hanging drops of Ras–Sos complex (2.5 μ l,

10 mg ml⁻¹) in buffer A were mixed with an equal amount of reservoir buffer containing 2.7–3.2 M sodium formate and 100 mM Tris buffer, pH 8.0, and kept at 4 °C. Crystal showers appeared after 1–2 days with large single crystals growing to full size (0.3 × 0.3 × 0.15 mm³) within 2–3 weeks. The crystals contain one heterodimeric complex per asymmetric unit and belong to space group *I*422 (*a* = *b* = 142.7 Å, *c* = 207.9 Å). Crystals were collected in 3.5 M sodium formate and 100 mM Tris buffer, pH 8.0, and cryoprotected in harvesting buffer with 10% (w/v) sucrose and 10% (v/v) ethylene glycol before flash freezing in liquid propane. Heavy-atom derivatives were prepared by soaking crystals in harvesting buffer containing the following heavy-atom solutions (soak durations were 1 day unless indicated otherwise): ethyl mercurithiosalicylate (EMTS), Baker's dimercurial, platinum terpyridine (12 h), 0.1 mM; osmium chloride (2 days), *p*-chloromercuribenzoic acid (PCMB), 1 mM; gold cyanide (2 days), 5 mM; trimethyl lead acetate (2 days), 10 mM; Table 1. All but one of the data sets used in this analysis were measured at Brookhaven National Laboratory on beamline X25 using the Brandeis 2 × 2 (four module) CCD-based detector³⁴. A mercury derivative data set (PCMB) was measured at Cornell High Energy Synchrotron Source on beamline F2 using the Q1 CCD-based detector (ADSC). Data processing was performed using Denzo, and data reduction was performed using Scalepack³⁵. MIR phases were calculated using MLPHARE as implemented in the CCP4 suite of programs³⁶. Solvent flattening was performed using DM³⁷.

Model building and refinement. Model building was performed using O³⁸. Coordinates for Ras bound to a GTP analogue (5P21)³⁹ were obtained from the Protein Data Bank⁴⁰, and the molecule, with Switch 1 and Switch 2 regions deleted, was fit into density. After an initial round of model building and positional refinement using CNS⁴¹ with bulk solvent corrections and anisotropic *B*-factor scaling protocols, phase combination methods using Sigma⁴² resulted in a much improved map into which the Switch 2 region of Ras, the entire catalytic domain and the N-terminal helices of Sos were built. Electron density maps based on multiple simulated annealing models⁴³ allowed the remaining regions of Ras and Sos to be placed into density. Residues 564–567 (N-terminal), 591–597, 654–675, 742–751 and 1045–1049 (C-terminal) are disordered and not modelled in Sos; no residues of Ras are disordered. The Ramachandran plot shows that 89% of all residues are in the most favoured regions and no residues are in disallowed regions.

Received 6 May; accepted 10 June 1998.

- Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117–127 (1991).
- Boguski, M. S. & McCormick, F. Proteins regulating Ras and its relatives. *Nature* **366**, 643–654 (1993).
- Medema, R. H., de Vries-Smits, A. M., van der Zon, G. C. M., Maassen, J. A. & Bos, J. L. Ras activation by insulin and epidermal growth factor through enhanced exchange of guanine nucleotides on p21ras. *Mol. Cell. Biol.* **13**, 155–162 (1993).
- Buday, L. & Downward, J. Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell* **73**, 611–620 (1993).
- Gale, N. W., Kaplan, S., Lowenstein, E. J., Schlessinger, J. & Bar-Sagi, D. Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on Ras. *Nature* **363**, 88–92 (1993).
- Bar-Sagi, D. The Sos (Son of sevenless) protein. *Trends Endocrin. Metab.* **5**, 165–169 (1994).
- Schlessinger, J. How receptor tyrosine kinases activate Ras. *Trends Biochem. Sci.* **18**, 273–275 (1994).
- Yu, H. & Schreiber, S. L. Structure of guanine-nucleotide-exchange factor human Mss4 and identification of its Rab-interacting surface. *Nature* **376**, 788–791 (1995).
- Mossessova, E., Gulbis, J. M. & Goldberg, J. Structure of the guanine nucleotide exchange factor Sec7 domain of human Arno and analysis of the interaction with ARF GTPase. *Cell* **92**, 415–423 (1998).
- Cherfils, J. *et al.* Structure of the Sec7 domain of the Arf exchange factor ARNO. *Nature* **392**, 101–105 (1998).
- Renault, L. *et al.* The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature* **392**, 97–101 (1998).
- Wang, Y., Jiang, Y., Meyering-Voss, M., Sprinzl, M. & Sigler, P. B. Crystal structure of the EF-Tu-EF-Ts complex from *Thermus thermophilus*. *Nat. Struct. Biol.* **4**, 650–656 (1997).
- Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. & Leberman, R. The structure of the *Escherichia coli* EF-Tu-EF-Ts complex at 2.5 Å resolution. *Nature* **379**, 511–518 (1996).
- Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).

- Chardin, P. *et al.* Human Sos 1: A guanine nucleotide exchange factor for Ras that binds to GRB2. *Science* **260**, 1338–1343 (1993).
- Lenzen, C., Cool, R. H., Prinz, H., Kuhlmann, J. & Wittinghofer, A. Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25⁴⁴. *Biochemistry* **37**, 7420–7430 (1998).
- Lai, C.-C., Boguski, M., Broek, D. & Powers, S. Influence of guanine nucleotides on complex formation between Ras and Cdc25 proteins. *Mol. Cell. Biol.* **13**, 1345–1352 (1993).
- Mistou, M. Y. *et al.* Mutations of H-Ras p21 that define important regions for the molecular mechanism of the CDC25 C-domain, a guanine nucleotide dissociation stimulator. *EMBO J.* **11**, 2391–2397 (1992).
- Milburn, M. V. *et al.* Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247**, 939–945 (1990).
- Jurnak, F. Structure of the GDP domain of EF-Tu and location of the amino acids homologous to ras oncogene proteins. *Science* **230**, 32–36 (1985).
- Powers, S., O'Neill, K. & Wigler, M. Dominant yeast and mammalian Ras mutants that interfere with CDC25-dependent activation of wild-type Ras in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9**, 390–395 (1989).
- Haney, S. A. & Broach, J. R. Cdc25p, the guanine nucleotide exchange factor for the Ras proteins of *Saccharomyces cerevisiae*, promotes exchange by stabilizing ras in a nucleotide free state. *J. Biol. Chem.* **269**, 16541–16548 (1994).
- Klebe, C., Prinz, H., Wittinghofer, A. & Goody, R. S. The kinetic mechanism of Ran-nucleotide exchange catalyzed by RCC1. *Biochemistry* **34**, 12543–12552 (1995).
- Verrotti, A. C. *et al.* Ras residues that are distant from the GDP binding site play a critical role in dissociation factor-stimulated release of GDP. *EMBO J.* **11**, 2855–2862 (1992).
- Segal, M., Willumsen, B. M. & Levitzki, A. Residues crucial for Ras interaction with GDP-GTP exchangers. *Proc. Natl Acad. Sci. USA* **90**, 5564–5568 (1993).
- Mosteller, R. D., Han, J. & Broek, D. Identification of residues of the H-Ras protein critical for functional interaction with guanine nucleotide exchange factors. *Mol. Cell. Biol.* **14**, 1104–1112 (1994).
- Segal, M., Marbach, I., Willumsen, B. M. & Levitzki, A. Two distinct regions of Ras participate in functional interaction with GDP-GTP exchangers. *Eur. J. Biochem.* **228**, 96–101 (1995).
- Leonardsen, L., DeClue, J. E., Lybaek, H., Lowy, D. R. & Willumsen, B. M. Ras p21 sequences opposite the nucleotide binding pocket are required for GRF-mediated nucleotide release. *Oncogene* **13**, 2177–2187 (1996).
- Crech, J. B., Bernardi, A. & Parmeggiani, A. Distal switch II region of Ras2p is required for interaction with guanine nucleotide exchange factor. *J. Biol. Chem.* **271**, 17234–17240 (1996).
- Quilliam, L. A. *et al.* Involvement of the switch 2 domain of Ras in its interaction with guanine nucleotide exchange factors. *J. Biol. Chem.* **271**, 11076–11082 (1996).
- Feig, L. A. & Cooper, G. M. Inhibition of NIH 3T3 cell proliferation by a mutant Ras protein with preferential affinity for GDP. *Mol. Cell. Biol.* **8**, 3235–3243 (1988).
- Chen, S.-Y., Huff, S. Y., Lai, C.-C., Der, C. J. & Powers, S. Ras-15A protein shares highly similar dominant-negative biological properties with Ras-17N and forms a stable, guanine-nucleotide resistant complex with CDC25 exchange factor. *Oncogene* **9**, 2691–2698 (1994).
- Powers, S., Gonzales, E., Christensen, T., Cubert, J. & Broek, D. Functional cloning of Bud5, a Cdc25-related gene from *S. cerevisiae* that can suppress a dominant-negative Ras2 mutant. *Cell* **65**, 1225–1231 (1991).
- Westbrook, E. M. & Naday, I. Charge-coupled device-based area detectors. *Methods Enzymol.* **276**, 244–268 (1997).
- Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Quilliam, L. A. *et al.* ESRF-EACBM Newsletter. *Protein Crystallogr.* **31**, 34–38 (1994).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Pai, E. F. *et al.* Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **9**, 2351–2359 (1990).
- Bernstein, F. C. *et al.* The protein data bank: a computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.* **185**, 584–591 (1978).
- Brünger, A. T. *et al.* Crystallography and NMR system (CNS): A new software suite for macromolecular structure determination. *Acta Crystallogr. D* (in press).
- Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).
- Brünger, A. T., Adams, P. D. & Rice, L. M. New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure* **5**, 325–336 (1997).
- Carson, M. Ribbons 2.0. *J. Appl. Crystallogr.* **24**, 958–961 (1991).
- Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).
- Pai, E. F. *et al.* Structure of the guanine-nucleotide-binding domain of Ha-ras oncogene product p21 in the triphosphate conformation. *Nature* **341**, 209–214 (1989).

Acknowledgements. We thank L. Leighton and M. Uy for technical assistance; J. Bonanno, S. K. Burley, R. H. Chen, X. Chen, E. Conti, S. Corbalan-Garcia, D. Galron, B. Hall, M. Huse, S. Jacques, D. Jeruzalmi, C.-H. Lee, S. Soisson, H. Yamaguchi, S.-S. Yang and Y. Zhang for advice and help; L. Berman, R. Sweet, the staff of NSLS and CHESS for help with synchrotron data collection; and A. Wittinghofer and R. H. Cool for sharing a manuscript prior to publication. P.A.B.-S. is supported by grants from the Cystic Fibrosis Foundation and the NIH. D.B.-S. acknowledges grant support from the NIH.

Correspondence and requests for materials should be addressed to J.K. The atomic coordinates have been deposited in the Brookhaven Protein Data Bank under code 1BKD.

A millisecond pulsar in an X-ray binary system

Rudy Wijnands* & Michiel van der Klis*†

*Astronomical Institute "Anton Pannekoek", University of Amsterdam, and Center for High Energy Astrophysics, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands

†Department of Astronomy, University of California, Berkeley, Berkeley, California 94720, USA

Ordinary radio pulsars¹ are neutron stars with magnetic fields of $\sim 10^{12}$ gauss and spin periods in the range 0.1 to 3 seconds. In contrast, millisecond radio pulsars² have much weaker fields ($\sim 10^9$ gauss) and faster, millisecond spin rates. For both types of pulsar, the energy driving the radio pulsations is thought to be derived from the rotation of the neutron star. The star gradually 'spins down' as energy is radiated away. Millisecond radio pulsars are often located in binary systems³. In a widely accepted theoretical model^{4,5}, they started as ordinary pulsars which lost most of their magnetic field and were 'spun up' to millisecond periods by the accretion of matter from a companion star in an X-ray binary system. Evidence^{6–11} for this model has gradually mounted, but direct proof—in the form of the predicted coherent millisecond X-ray pulsations in the persistent flux of an X-ray binary has been lacking, despite many searches^{12–15}. Here we report the discovery¹⁶ of such a pulsar, confirming theoretical expectations. The source will probably become a millisecond radio pulsar when the accretion turns off completely.

The transient X-ray burst source SAX J1808.4–3658 was discovered¹⁷ in September 1996 with the Wide Field Cameras on board BeppoSAX. Two X-ray bursts of type I, generally accepted to be due to a thermonuclear runaway on the surface of a neutron star, were observed, and the source was classified as a low-mass X-ray binary (LMXB), containing a low-magnetic-field neutron star accreting matter from a companion star of less than one solar mass. The probable distance of the source is¹⁷ 4 kpc. Recently, a transient X-ray source designated XTE J1808–369 was detected¹⁸ with the Proportional Counter Array (PCA) on board the Rossi X-ray Timing Explorer (RXTE). This source is positionally coincident to within a few arcminutes with SAX J1808.4–3658, and it is likely that both sources are the same object.

XTE J1808–359 was observed by the RXTE team at Goddard Space Flight Center with the PCA on 11 April 1998. The data were placed in the online public RXTE data archive, allowing rapid analysis of the data by the scientific community. We obtained the data and performed a series of fast Fourier transforms designed to be sensitive to various types of rapid variability, including strictly

periodic pulsations¹⁹. We discovered¹⁶ clear evidence for X-ray pulsations with a frequency near 401.0 Hz in the resulting power-density spectrum (Fig. 1). The frequency was stable to within 0.01%, far exceeding the stability of any previously known phenomenon in an accreting low-magnetic-field neutron star, and making it practically certain that we were detecting the neutron-star spin rate. Triggered by this discovery, a series of further observations was performed²⁰. Here we report on the initial, 2-h RXTE observing span.

During this observation the pulsation frequency drifted slightly, due to the Doppler shifts caused by the orbital motion of the satellite around the Earth, and the pulsar around its companion star. Although the satellite ephemeris was not yet available at the time of our analysis and we could not therefore determine the pulsar's orbit, we were able to correct empirically for the frequency drift and to fold the pulsations coherently for several million cycles (Fig. 2). This leaves no other interpretation than the presence of a millisecond X-ray pulsar in the system¹⁶. The fact that this pulsar is detected in the (accretion-driven) outburst of a transient X-ray source shows that during our observations it was powered by accretion from a companion star, and not by its rotation.

In a subsequent analysis Chakrabarty and Morgan^{20,21} have found the orbital period of this pulsar (~ 2 h, close to the satellite orbital period) and measured its orbit. A detailed analysis of the pulsar's orbit is presented in this companion paper²⁰ so we shall not repeat it here. This work confirms that the pulsar is a member of a low-mass X-ray binary system, and provides important additional information on the stability of the pulsation frequency, which will eventually allow determination of whether the pulsar is still spinning up, or has already reached an equilibrium spin rate.

We can use the observed spin rate of the millisecond X-ray pulsar to constrain the magnetic field strength B of the neutron star, and thus test the prediction that it should be similar to that of millisecond radio pulsars, as would be required for this system to be the missing link between X-ray binaries and millisecond radio pulsars. For accretion to proceed, the centrifugal force on the accreting matter co-rotating in the magnetosphere must be less than the local gravitational force²². Using standard magnetospheric disk accretion theory²³ this leads to an upper limit on B of $(2-6) \times 10^8$ G, far below the $\sim 10^{12}$ G typical of slower accretion-powered pulsars. Centrifugal inhibition of the accretion occurs when the magnetospheric radius $r_M = 18 \xi \mu_{26}^{4/7} m^{1/7} R_6^{-2/7} L_{37}^{-2/7}$ km is larger than the co-rotation radius $r_{co} = 15 m^{1/3} P_{\text{ms}}^{2/3}$ km, where μ_{26} is the magnetic moment in units of 10^{26} G cm³, L_{37} is the bolometric luminosity in units of 10^{37} erg s⁻¹, m the neutron-star mass in solar masses, R_6 its radius in units of 10^6 cm, ξ a dimensionless factor that is probably between 0.5 and 1, and P_{ms} is the spin period in ms. Given the 2.49-ms spin period and a bolometric luminosity of $\sim 6 \times 10^{36}$ erg s⁻¹, estimated from the information available on the X-ray spectrum^{17,18,24}, this can be used to set an upper limit of 31 km on r_M and thus to constrain B .

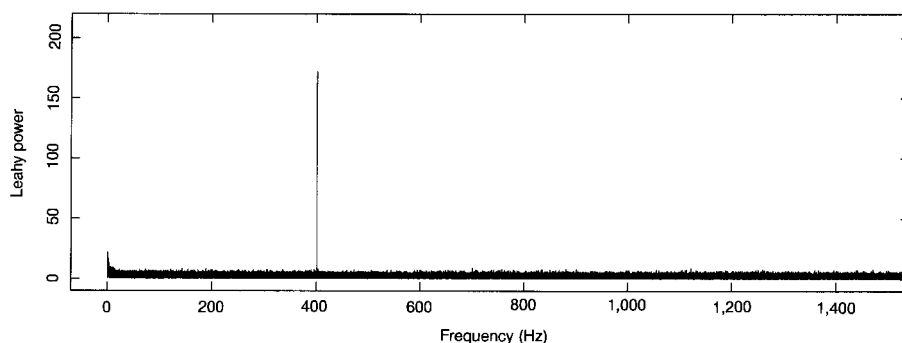


Figure 1 Leahy¹² normalized power-density spectrum of the 1998 April 11 20:38–21:21 UT persistent emission of XTE J1808–369. The 2–60 keV source count rate varied between 892 and 1007 counts s⁻¹, including a ~ 130 counts s⁻¹ background.

No X-ray bursts were observed. The spectrum shows clear evidence for X-ray pulsations near a frequency of 401.0 Hz, with an amplitude of $\sim 4.1\%$ (r.m.s.) of the source flux in the 2–60 keV range.

In calculating the upper limit on B given above, we have assumed a $1.4M_{\odot}$ neutron star with a radius of 10 km. As the bolometric luminosity drops, the limit on B comes down further.

Such a low magnetic field strength is indeed in the range of values inferred for millisecond radio pulsars²⁵. The object seems to be located above the radio pulsar "death line" (below which the field is too weak for the radio pulsar mechanism to operate²⁵), so when the accretion phase of the binary finally ends and the source turns off as an accretion-powered X-ray pulsar, it would probably switch on as a rotation-powered radio pulsar. This might even happen between X-ray outbursts²⁶, if the density of the accreting matter in the pulsar environment, at least out to the light cylinder at 120 km, drops sufficiently. So, this LMXB could indeed be one of the progenitors of the millisecond radio pulsars, and it might be an intermittent radio pulsar now.

We do not know the circumstances under which the neutron star was spun up to its present rapid rate. Current data^{17,18} suggest a typical outburst mass accretion rate of $5 \times 10^{-10} M_{\odot} \text{ yr}^{-1}$ and a duty cycle of $\sim 5\%$ for the transient outbursts. At that rate, the total spin up time would be $\sim 2 \times 10^9 \text{ yr}$.

Previous indications for high spin rates and low field strengths of the neutron stars in LMXBs have been numerous, but were all indirect. Evidence for high (several hundred Hz) spin rates came from the detection⁶⁻⁸ of frequencies interpreted⁸⁻¹¹ as beat frequencies between the star's spin frequency and the orbital frequency of clumps of matter orbiting the star. Further evidence for this was obtained from drifting frequencies observed⁸ during thermonuclear X-ray bursts, which are probably caused by thermonuclear hotspots in the star's surface layers, spinning around at approximately the star's spin rate. Indications of low (10^8 to several 10^9 G) field

strengths came from thermonuclear burst theory²⁷⁻²⁹, which says that the occurrence of X-ray bursts requires the field to be weak, and from various other indirect methods^{9,10,30,31}. The spin and field we deduce for XTE J1808-369 are in accordance with these previous estimates for other LMXBs. If XTE J1808-369 is the same source as SAX J1808.4-3658, then the source is the first pulsar to show both coherent pulsations in its persistent emission and thermonuclear bursts, in agreement with our conclusion that XTE J1808-369 is by far the lowest-field accretion-driven pulsar known.

As both the spin rate and the field strength of XTE J1808-369 resemble those inferred for other, non-pulsing LMXBs, the question arises: why is XTE J1808-369 the only known LMXB with a millisecond X-ray pulsar? The pulsations in XTE J1808-369 are strong enough to ensure that similar-amplitude pulsations would have been noticed in other systems as well. The most straightforward explanation seems to be that the magnetic field of XTE J1808-369 is considerably stronger than that of other systems of similar (relatively low) X-ray luminosity. The more efficient magnetic channelling of the accreting matter in XTE J1808-369 would then lead to higher pulsation amplitudes, making the pulsar easier to detect. This would mean that the fields of the other low-luminosity LMXBs have been overestimated, which is possible in view of the large uncertainties in those estimates. If this is true, then it remains to be seen whether these other systems could also be millisecond radio pulsar progenitors such as we deduce XTE J1808-369 to be. As XTE J1808-369 is a weak, transient X-ray source, it is possible that it is the first example of a class of millisecond radio pulsar progenitors that has not been noticed until now. $\square$

Received 21 April; accepted 16 June 1998.

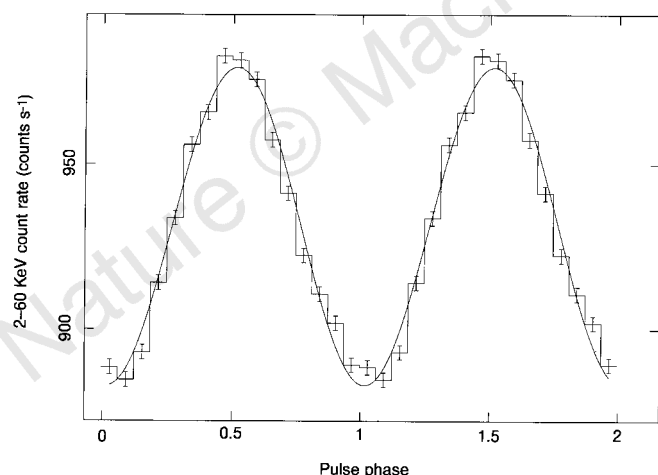


Figure 2 The 2-60 keV light curve folded at the 2.49-ms period. Two cycles are plotted for clarity. The solid line is the best-fitting sinusoid. To produce this plot, we divided the data into 128-s segments and determined the phase and the frequency of the pulsations in each segment by an epoch-folding technique. The results were consistent with a smooth, approximately sinusoidal ~ 110 -min variation of the frequency between 400.986 and 401.022 Hz. After correcting the phases for the effects of this drift, the pulsations could be folded coherently. The resulting grand-average folded pulse profile is near-sinusoidal, with no significant difference between different photon energy bands. Fitting the folded light curve with a harmonic series, we find that the amplitudes of the harmonics at 2 and 3 times the 401.0-Hz frequency are $\sim 11\%$ and $\sim 4\%$ of the 401.0-Hz amplitude, respectively. There is no evidence that the true spin frequency is 200.5 Hz: we find no significant difference in shape between the odd and even pulses when folding the data on twice the period and harmonics at 1, 3 and 5 times 200.5 Hz are not detected, with typical upper limits on the amplitude of $\sim 5\%$ of the 401.0 Hz amplitude.

- Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F. & Collins, R. A. Observation of a rapidly pulsating radio source. *Nature* **217**, 709-713 (1968).
- Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M. & Goss, W. M. A millisecond pulsar. *Nature* **300**, 615-618 (1982).
- Hulse, R. A. & Taylor, J. H. Discovery of a pulsar in a binary system. *Astrophys. J.* **195**, L51-L53 (1975).
- Radhakrishnan, V. & Srinivasan, G. On the origin of the recently discovered ultra-rapid pulsar. *Curr. Sci.* **51**, 1096-1099 (1982).
- Alpar, M. A., Cheng, A. F., Ruderman, M. A. & Shaham, J. A new class of radio pulsars. *Nature* **300**, 728-730 (1982).
- van der Klis, M. *et al.* Intensity-dependent quasi-periodic oscillations in the X-ray flux of GX 5-1. *Nature* **316**, 225-230 (1985).
- van der Klis, M. *et al.* Discovery of submillisecond quasi-periodic oscillations in the X-ray flux of Scorpius X-1. *Astrophys. J.* **469**, L1-L4 (1996).
- Strohmayer, T. E. *et al.* Millisecond X-ray variability from an accreting neutron star system. *Astrophys. J.* **469**, L9-L12 (1996).
- Alpar, M. A. & Shaham, J. Is GX 5-1 a millisecond pulsar? *Nature* **316**, 239-241 (1985).
- Lamb, F. K., Shibasaki, N., Alpar, M. A. & Shaham, J. Quasi-periodic X-ray oscillations in bright galactic bulge X-ray sources. *Nature* **317**, 681-687 (1985).
- Miller, M. C., Lamb, F. K. & Psaltis, D. Sonic-point model of kilohertz QPOs in LMXBs. *Astrophys. J.* (in the press).
- Leahy, D. A. *et al.* On searches for pulsed emission with application to four globular cluster X-ray sources: NGC 1851, 6441, 6624 and 6712. *Astrophys. J.* **266**, 160-170 (1983).
- Mereghetti, S. & Grindlay, J. E. A search for millisecond periodic and quasi-periodic pulsations in low-mass X-ray binaries. *Astrophys. J.* **312**, 727-731 (1987).
- Wood, K. S. *et al.* Searches for millisecond pulsations in low-mass X-ray binaries. *Astrophys. J.* **379**, 295-309 (1991).
- Vaughan, B. A. *et al.* Searches for millisecond pulsations in low-mass X-ray binaries. II. *Astrophys. J.* **435**, 362-369 (1994).
- Wijnands, R. & van der Klis, M. SAX J1808.4-3658 = XTE J1808-369. *IAU Circ.* No. 6876 (1998).
- in 't Zand, J. J. M. *et al.* Discovery of the X-ray transient XAX J1808.4-3658, a likely low-mass X-ray binary. *Astron. Astrophys.* **331**, L25-L28 (1998).
- Marshall, F. E. SAX J1808.4-3658 = XTE J1808-369. *IAU Circ.* No. 6876 (1998).
- van der Klis, M. Fourier techniques in X-ray timing. In *Timing Neutron Stars* (eds Ögelman, H. & van den Heuvel, E. P. J.) 27-69 (Kluwer Academic, 1989).
- Chakrabarty, D. & Morgan, E. H. The two-hour orbit of a binary millisecond X-ray pulsar. *Nature* **394**, 346-348 (1998).
- Chakrabarty, D. & Morgan, E. H. SAX J1808.4-3658 = XTE J1808-369. *IAU Circ.* No. 6877 (1998).
- Stella, L., White, N. E. & Rosner, R. Intermittent stellar wind accretion and the long-term activity of population I binary systems containing an X-ray pulsar. *Astrophys. J.* **308**, 669-679 (1986).
- Ghosh, P. & Lamb, F. K. Accretion by rotating magnetic neutron stars. II. Radial and vertical structure of the transition zone in disk accretion. *Astrophys. J.* **234**, L296-L316 (1979).
- Heindl, W., Marsden, D. & Blanco, P. SAX J1808.4-3658 = XTE J1808-369. *IAU Circ.* No. 6878 (1998).
- Bhattacharya, D. & van den Heuvel, E. P. J. Formation and evolution of binary and millisecond radio pulsars. *Phys. Rep.* **203**, 1-124 (1991).
- Campana, S. *et al.* Aquila X-1 from outburst to quiescence: the onset of the propeller effect and signs of a turned-on rotation-powered pulsar. *Astrophys. J.* **499**, L65-L68 (1998).
- Joss, P. C. & Li, F. K. Helium-burning flashes on accreting neutron stars: Effects of stellar mass radius, and magnetic field. *Astrophys. J.* **238**, 287-295 (1980).
- Bildsten, L. Propagation of nuclear burning fronts on accreting neutron stars: X-ray bursts and sub-hertz noise. *Astrophys. J.* **438**, 852-875 (1995).
- Bildsten, L. & Brown, E. F. Thermonuclear burning on the accreting X-ray pulsar GRO J1744-28. *Astrophys. J.* **477**, 897-904 (1997).

30. Psaltis, D., Lamb, F. K. & Miller, G. S. X-ray spectra of Z sources. *Astrophys. J.* **454**, L137–L140 (1995).
 31. White, N. E. & Zhang, W. Millisecond X-ray pulsars in low-mass X-ray binaries. *Astrophys. J.* **490**, L87–L90 (1997).

Acknowledgements. We thank L. Bildsten, J. van Paradijs, R. Fender and R. Ramachandran for comments on versions of this manuscript, the aforementioned, J. Arons and F. Verbunt for discussions, and the RXTE team for promptly making the data available online. M.K. was supported by the Visiting Miller Professor Program of the Miller Institute for Basic Research in Science (UCB). This work was supported by the Netherlands Organization for Research in Astronomy ASTRON.

Correspondence and requests for materials should be addressed to M.v.d.K. at the University of Amsterdam (e-mail: michiel@astro.uva.nl).

The two-hour orbit of a binary millisecond X-ray pulsar

Deepto Chakrabarty & Edward H. Morgan

Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Typical radio pulsars are magnetized neutron stars that are born rapidly rotating and slow down as they age on time scales of 10 to 100 million years. In contrast, millisecond radio pulsars spin very rapidly even though many are billions of years old¹. The most compelling explanation is that they have been ‘spun up’ by the transfer of angular momentum during the accretion of material from a companion star in so-called low-mass X-ray binary systems, LMXBs. (LMXBs consist of a neutron star or black hole accreting matter from a companion with mass less than one solar mass².) The recent detection of coherent X-ray pulsations with a millisecond period from a suspected low-mass X-ray binary system appears to confirm this link³. Here we report observations showing that the orbital period of this binary system is two hours, which establishes it as an LMXB. We also find an apparent modulation of the X-ray flux at the orbital period (at the two per cent level), with a broad minimum when the pulsar is behind the low-mass companion star. This system seems closely related to the ‘black-widow’ millisecond radio pulsars, which are evaporating their companions through irradiation^{4–8}. It may appear as an eclipsing radio pulsar during periods of X-ray quiescence.

The transient X-ray source SAX J1808.4–3658 was first observed in September 1996 by the Wide Field Cameras on the BeppoSAX X-ray satellite during a bright state which lasted about 20 days (ref. 9). The source was not detected (X-ray flux $< 3 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the 2–10 keV band) during an August 1996 observation, but reached a peak X-ray flux of $2 \times 10^{-9} \text{ erg cm}^{-2} \text{ s}^{-1}$ during the September bright state. Also during the bright state, BeppoSAX detected two type-I X-ray bursts from the source, each lasting less than 30 s. Such bursts are due to thermonuclear ignition of accreted material on the surface of neutron stars that have low ($< 10^{10} \text{ G}$) magnetic fields¹⁰. Analysis of the bursts in SAX J1808.4–3658 indicates that it is 4 kpc distant and has a peak X-ray luminosity of $6 \times 10^{36} \text{ erg s}^{-1}$ in its bright state and $< 10^{35} \text{ erg s}^{-1}$ in quiescence⁹. The source position was determined to within a 1σ radius of 0.6 arcmin.

More recently, a serendipitous slew of the Proportional Counter Array (PCA) on the Rossi X-Ray Timing Explorer (RXTE) over this region of the sky on 9 April 1998 indicated the presence of an X-ray source, designated XTE J1808–369. Repeated scans over the region with the PCA on 11 April localized the position with sufficient precision to confirm that the source was probably the same as SAX J1808.4–3658 (ref. 11). RXTE subsequently made numerous pointed observations of the source between 11 April and 6 May on a public ‘target-of-opportunity’ basis. The source reached a peak luminosity of $5 \times 10^{36} \text{ erg s}^{-1}$ on 13 April and had faded below $10^{35} \text{ erg s}^{-1}$ by 6 May.

Strong coherent 400-Hz pulsations were clearly detected in the 2–30 keV PCA data from most of these observations, and were first

reported by Wijnands and van der Klis^{3,12}. These are the fastest persistent coherent pulsations ever detected from an X-ray binary. From standard magnetic disk accretion theory, the detection of X-ray pulsations at a luminosity $\sim 10^{36} \text{ erg s}^{-1}$ requires that the neutron star’s surface dipole magnetic field be $B < 10^8 \text{ G}$ (ref. 3). This is much weaker than the $\sim 10^{12} \text{ G}$ fields found in other (slower) accretion-powered pulsars¹³, but it is consistent with the weak fields expected for type-I X-ray bursters¹⁰. If pulsations are detected when the source is at lower luminosities, the implied surface field would be even weaker.

In order to search for orbital Doppler shifts, we first selected the 3–30 keV photon arrival times during 11–18 April and binned them into 2^{-11}-s ($\sim 0.5\text{-ms}$) samples, after correcting these times for RXTE’s motion with respect to the Solar System barycentre using the best BeppoSAX position for the source⁹ (right ascension 18 h 08 min 29 s, declination $-36^\circ 58.6'$, equinox J2000.0). We then measured the barycentric pulse frequency at 128-s intervals using an oversampled Fourier power spectrum. The sequence of pulse frequencies showed an obvious 2-h sinusoidal modulation which was well fitted by a constant spin frequency plus a circular, keplerian orbit¹⁴.

Using this provisional timing model, we epoch-folded each 128-s interval in the 11–18 April data and cross-correlated the resulting pulse profiles with a sinusoid to measure the pulse phase history, which generally yields more precise timing parameters than a pulse frequency history. These pulse phases were well fitted (reduced $\chi^2 = 1.26$ with 441 degrees of freedom) by a constant spin frequency ν plus a circular keplerian orbit. The best-fit parameters are given in Table 1. The pulse arrival time delays with respect to a constant frequency model are plotted in Fig. 1. No improvement in the fit to the 11–18 April data is obtained by including an orbital eccentricity e or a spin frequency derivative $\dot{\nu}$. A more sensitive search for a non-zero value of $\dot{\nu}$, which includes the later data taken as the X-ray outburst faded from view, will be presented elsewhere.

The present limits on $\dot{\nu}$ constrain the positional uncertainty of the star. An error of 0.6 arcmin in the assumed source position would give rise to periodic timing residuals with 1-yr period and $\sim 80\text{-ms}$ amplitude¹⁵, and such residuals would lead to an apparent $\dot{\nu}$ in the timing solution for observations shorter than 3 months. Our limit on $\dot{\nu}$ from the 11–18 April data suggests that the BeppoSAX position is accurate to within 25 arcsec in ecliptic longitude. We note that a probable optical counterpart has been identified, lying 19 arcsec from the BeppoSAX position^{16,17}.

The pulsar mass function f_1 , which relates the pulsar mass m_1 , the companion mass m_2 , and the binary inclination i (the angle between the line of sight and the orbital angular momentum vector), may be computed from the observed keplerian parameters $a_1 \sin i$ and P_{orb} :

$$f_1 \equiv \frac{(m_2 \sin i)^3}{(m_1 + m_2)^2} = \frac{4\pi^2 (a_1 \sin i)^3}{GP_{\text{orb}}^2} \quad (1)$$

from which we find $f_1 = 3.7789(2) \times 10^{-5} M_\odot$ (here $M_\odot$ is the solar mass). Given assumed values of m_1 and i , equation (1) can be solved for m_2 . For m_1 , dynamical mass measurements in both binary millisecond radio pulsars¹⁸ and accretion-powered pulsars¹⁹ are all

Table 1 Observed parameters of SAX J1808.4–3658

Barycentric pulse frequency, ν_0 (Hz)	400.9752106(8)
Pulse frequency derivative, $ \dot{\nu} $ (Hz s^{-1})	$< 7 \times 10^{-13}$
Projected semimajor axis, $a_1 \sin i$ (light ms)	62.809(1)
Orbital period, P_{orb} (s)	7249.119(1)
Epoch of 90° mean longitude, $T_{\pi/2}$ (MJD)	50914.899440(1)
Eccentricity, e	$< 5 \times 10^{-4}$
Pulsar mass function, f_1 ($M_\odot$)	$3.7789(2) \times 10^{-5}$

Modified Julian date MJD = JD – 2400000.5. Numbers in parentheses are the 1σ uncertainties in the last significant figure, and upper limits are quoted at the 2σ level. The $\dot{\nu}$ limit refers to the magnitude, independent of sign. Epochs are given in Barycentric Dynamical Time (TDB).

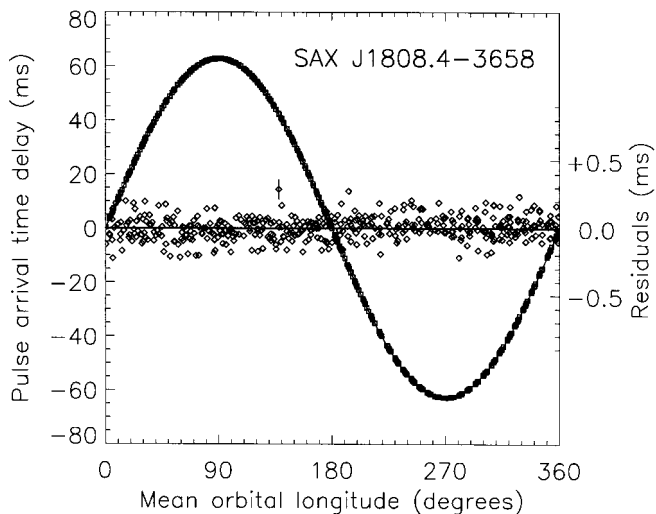


Figure 1 Pulse arrival time delays due to the 2-h binary orbit, with respect to a constant pulse frequency model. Mean longitude is the orbital phase angle measured from the ascending node, so that the pulsar is behind the companion at 90°. Squares show measured time delays, the solid curve is the best-fit orbital model, and diamonds show the model residuals (on a 50 $\times$ expanded scale). The residuals have an r.m.s. uncertainty of 75 μ s, shown for one of the points.

generally consistent with a neutron star mass of $1.35 M_{\odot}$. However, neutron stars in LMXBs have undergone substantial accretion from their binary companion and may be as massive as $2 M_{\odot}$ (see, for example, ref. 20). For i , we note that for an ensemble of binaries whose inclinations are distributed randomly, the *a priori* probability of observing a system with inclination i or smaller is $(1 - \cos i)$. If SAX J1808.4–3658 is drawn from such an ensemble, then there is a 95% probability that $m_2 < 0.14 M_{\odot}$ if $m_1 = 1.35 M_{\odot}$. For $m_1 = 2 M_{\odot}$, we find $m_2 < 0.18 M_{\odot}$. In either case, the binary separation is probably of the order of 1 light s.

The nature of the companion star provides an important clue to the evolutionary history of an LMXB. The only direct information we have on the companion comes from observations of the probable optical counterpart, which suggest that the companion is a faint low-mass star subject to X-ray heating by the pulsar^{16,17}. We can use the binary parameters to deduce more about this star. The companions in most LMXBs fill (or nearly fill) their critical gravitational potential surface, known as the Roche lobe. Because the mean density of a Roche-lobe-filling companion (for the case where $m_2 < m_1$) is uniquely determined by the binary period²¹, we can relate the companion's radius in SAX J1808.4–3658 to its mass as $R_2 = 0.17 R_{\odot} (m_2/0.1 M_{\odot})^{1/3}$, giving a minimum companion radius of $0.12 R_{\odot}$ (here $R_{\odot}$ is the solar radius). The assumption of Roche-lobe overflow thus strongly constrains the nature of the companion. Given the compact size of the binary, only white-dwarf or low-mass main-sequence companions are plausible. (In principle, a helium-burning star could also fit in this binary, but we would expect it to be much brighter than the proposed optical counterpart²².) A white dwarf with mass m has radius $R = 0.013 R_{\odot} (1 + X)^{5/3} (m/M_{\odot})^{-1/3}$, where X is the hydrogen mass fraction²³. Comparing this to our equation for R_2 , we see that a helium white dwarf ($X = 0$) would be too small, and hence is ruled out. Even a hydrogen-rich white dwarf ($X = 0.9$) is essentially excluded for all but $i \approx 90^\circ$. However, the lack of a deep X-ray eclipse rules out $i > 82^\circ$, assuming a Roche-lobe-filling companion.

Normal low-mass hydrogen main-sequence stars have $R/R_{\odot} \approx m/M_{\odot}$, and comparison of detailed models²⁴ with our equation for R_2 yields a Roche-lobe-filling mass of $0.17 M_{\odot}$. This would require a small binary inclination ($i < 20^\circ$), which has low *a*

priori probability (5%). However, normal main-sequence models are inappropriate in our case, as the stellar structure must be strongly influenced by irradiation from the pulsar. If the incident flux at the companion surface exceeds $\sim 10^{10} \text{ erg cm}^{-2} \text{ s}^{-1}$, the star may be “bloated” to larger radii²⁵. This is especially true for main-sequence stars having $m < 0.3 M_{\odot}$, which have fully convective envelopes. For the binary separation in this system ~ 1 light s, the X-ray flux at the companion will exceed $10^{14} \text{ erg cm}^{-2} \text{ s}^{-1}$, and the bloating could be severe. Under these conditions, main-sequence stars as light as $0.1 M_{\odot}$ or less may fill their Roche lobe, consistent with a more likely inclination.

Another possibility is that irradiation by the neutron star might directly drive mass loss from the companion (‘ablation’) whether or not it fills its Roche lobe²⁶. The binary parameters of SAX J1808.4–3658 are very similar to those of the five “black widow” millisecond radio pulsars in very close binaries, all of which are ablating their low-mass companions^{4–8}. At least four of these radio systems show eclipses which are too broad to be caused by a Roche-lobe-filling companion, and which are instead attributed to an ablated wind^{4–6,8}.

We find evidence that the X-ray flux from SAX J1808.4–3658 is slightly modulated at the orbital period. The apparent modulation is roughly sinusoidal with 2% amplitude and a minimum when the pulsar is behind the companion, and it is much broader and much shallower than the (at most) 5-min eclipse possible from a Roche-lobe-filling companion. However, the detailed features of this 2-h modulation must be viewed with caution, as its strength is comparable to the variation of the RXTE background with the 96-min spacecraft orbit. The similarity in strength and timescales of these two variations makes them difficult to disentangle, and final confirmation must await either additional data or an improvement in the RXTE background model. Even so, the detection of a 2-h source flux modulation (if not its precise morphology) seems secure.

We suggest that the intensity dips are due to scattering in an ablated wind. Given the similarity between this source and the eclipsing radio pulsars, SAX J1808.4–3658 may emerge as a radio pulsar during X-ray quiescence²⁷. Moreover, because the radio emission would be less penetrating than the X-rays, a deeper eclipse dip might be observed, possibly providing a strong constraint on the binary inclination. The presence of the slight X-ray dips, if confirmed, would rule out small binary inclinations and make it highly likely that the companion mass is less than $0.1 M_{\odot}$.

Whether the mass accretion is fed by Roche-lobe overflow or an ablated wind or both, we can understand the transient nature of the X-ray emission as long as some of this material forms an accretion disk. Although the instantaneous mass accretion rate during the 1996 and 1998 bright states was $\sim 3 \times 10^{-10} M_{\odot} \text{ yr}^{-1}$, the small duty cycle of the X-ray activity indicates a long-term mean mass transfer rate of $\dot{M} \approx 1 \times 10^{-11} M_{\odot} \text{ yr}^{-1}$. For such a low mass transfer rate, the accretion disk would probably be subject to dwarf-nova-type instabilities, leading to episodic outbursts of X-ray emission^{28,29}. We note that for a Roche-lobe-filling companion, angular-momentum losses due to gravitational radiation would drive a mass transfer rate of $\sim 10^{-11} M_{\odot} \text{ yr}^{-1}$ for a $0.05 M_{\odot}$ secondary³⁰, consistent with our inferred inclination constraints.

Future X-ray outbursts from SAX J1808.4–3658 may allow detection of a spin frequency derivative, which would further constrain both the mass-accretion rate and the pulsar's magnetic field strength, and would provide a probe of magnetic disk accretion torque theory in the previously unexplored regime of a very small magnetosphere. The characteristic accretion torque expected for the observed X-ray luminosity should cause a spin frequency derivative of the order of $10^{-14} \text{ Hz s}^{-1}$, which is detectable in observations spanning a month or more. Further observations may also detect orbital period evolution, which would probe the competing effects of gravitational radiation, tidal interactions and mass loss on the binary angular momentum. An astrophysically interesting limit on

the mean orbital period derivative of $<10^{-11}$ could be reached in less than a year of observations. □

Received 29 April; accepted 16 June 1998.

- Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M. & Goss, W. M. A millisecond pulsar. *Nature* **300**, 615–618 (1982).
- Bhattacharya, D. & van den Heuvel, E. P. J. Formation and evolution of binary and millisecond pulsars. *Phys. Rep.* **203**, 1–124 (1991).
- Wijnands, R. & van der Klis, M. A millisecond pulsar in an X-ray binary system. *Nature* **394**, 344–346 (1998).
- Fruchter, A. S. *et al.* The eclipsing millisecond pulsar PSR 1957+20. *Astrophys. J.* **351**, 642–650 (1990).
- Nice, D. J. & Thorsett, S. E. Pulsar PSR 1744–24A: timing, eclipses, and the evolution of neutron star binaries. *Astrophys. J.* **397**, 249–259 (1992).
- Manchester, R. N. *et al.* Discovery of ten millisecond pulsars in the globular cluster 47 Tucanae. *Nature* **352**, 219–221 (1991).
- Deich, W. T. S. *et al.* The binary pulsar PSR 1908+00 in NGC 6760. *Astrophys. J.* **410**, L95–L98 (1993).
- Stappers, B. W. *et al.* Probing the eclipse region of a binary millisecond pulsar. *Astrophys. J.* **465**, L119–L122 (1996).
- in 't Zand, J. J. *et al.* Discovery of the X-ray transient SAX J1808.4–3658, a likely low-mass X-ray binary. *Astron. Astrophys.* **331**, L25–L28 (1998).
- Lewin, W. H. G., van Paradijs, J. & Taam, R. E. in *X-Ray Binaries* (eds Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J.) 175–232 (Cambridge Univ. Press, 1995).
- Marshall, F. E. SAX J1808.4–3658 = XTE J1808–369. *IAU Circ. No.* 6876 (1998).
- Wijnands, R. & van der Klis, M. SAX J1808.4–3658 = XTE J1808–369. *IAU Circ. No.* 6876 (1998).
- White, N. E., Nagase, F. & Parmar, A. N. in *X-Ray Binaries* (eds Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J.) 1–57 (Cambridge Univ. Press, 1995).
- Chakrabarty, D. & Morgan, E. H. SAX J1808.4–3658 = XTE J1808–369. *IAU Circ. No.* 6877 (1998).
- Lyne, A. G. & Graham-Smith, F. *Pulsar Astronomy* (Cambridge Univ. Press, 1990).
- Roche, P. *et al.* SAX J1808.4–3658 = XTE J1808–369. *IAU Circ. No.* 6885 (1998).
- Giles, A. B., Hill, K. M. & Greenhill, J. G. SAX J1808.4–3658 = XTE J1808–369. *IAU Circ. No.* 6886 (1998).
- Thorsett, S. E. & Chakrabarty, D. Neutron star mass measurements. I. Radio pulsars. *Astrophys. J.* (submitted); preprint astro-ph/9803260 at (<http://xxx.lanl.gov>) (1998).
- van Kerkwijk, M. H., van Paradijs, J. & Zuiderwijk, E. J. On the masses of neutron stars. *Astron. Astrophys.* **303**, 497–501 (1995).
- Zhang, W., Strohmayer, T. E. & Swank, J. H. Neutron star masses and radii as inferred from kilohertz quasi-periodic oscillations. *Astrophys. J.* **L167**–L170 (1997).
- Faulkner, J., Flannery, B. P. & Warner, B. Ultrashort-period binaries. II. HZ 29 (=AM CVn): a double-white-dwarf semidetached postcataclysmic nova? *Astrophys. J.* **175**, L79–L83 (1972).
- Savonije, G. J., de Kool, M. & van den Heuvel, E. P. J. The minimum orbital period for ultra-compact binaries with helium burning secondaries. *Astron. Astrophys.* **155**, 51–57 (1986).
- Paczynski, B. Gravitational waves and the evolution of close binaries. *Acta Astron.* **17**, 287–296 (1967).
- Tout, C. A., Pols, O. R., Eggleton, P. P. & Han, Z. Zero-age main-sequence radii and luminosities as analytic functions of mass and metallicity. *Mon. Not. R. Astron. Soc.* **281**, 257–262 (1996).
- Podsiadlowski, P. Irradiation-driven mass transfer in low-mass X-ray binaries. *Nature* **350**, 136–138 (1991).
- D'Antona, F. in *Evolutionary Processes in Binary Stars* (eds Wijers, R. A. M. J., Davies, M. B. & Tout, C. A.) 287–306 (Kluwer, Dordrecht, 1996).
- Stella, L., Campana, S., Colpi, M., Mereghetti, S. & Tavecchio, M. Do quiescent soft X-ray transients contain millisecond radio pulsars? *Astrophys. J.* **423**, L47–L50 (1994).
- van Paradijs, J. On the accretion instability in soft X-ray transients. *Astrophys. J.* **464**, L139–L141 (1996).
- King, A. R., Kolb, U. & Burderi, L. Black hole binaries and X-ray transients. *Astrophys. J.* **464**, L127–L130 (1996).
- Verbunt, F. & van den Heuvel, E. P. J. in *X-Ray Binaries* (eds Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J.) 457–494 (Cambridge Univ. Press, 1995).

Acknowledgements. We thank L. Bildsten, V. Kaspi, A. Levine, R. Nelson, R. Remillard, F. Rasio, S. Thorsett, M. van der Klis and B. Vaughan for discussions, and M. Muno for assistance with the data analysis. We also thank F. Marshall, J. Swank and the RXTE team at NASA/Goddard Space Flight Center for arranging these target-of-opportunity observations and the necessary follow-up. This work was supported by NASA.

Correspondence and requests for materials should be addressed to D.C. (e-mail: deepto@space.mit.edu).

Optical alignment and spinning of laser-trapped microscopic particles

M. E. J. Friese, T. A. Nieminen, N. R. Heckenberg & H. Rubinsztein-Dunlop

Centre for Laser Science, Department of Physics, The University of Queensland, Brisbane, Queensland 4072, Australia

Light-induced rotation of absorbing microscopic particles by transfer of angular momentum from light to the material raises the possibility of optically driven micromachines. The phenomenon has been observed using elliptically polarized laser beams¹ or beams with helical phase structure^{2,3}. But it is difficult to develop high power in such experiments because of overheating and unwanted axial forces, limiting the achievable rotation rates to a

few hertz. This problem can in principle be overcome by using transparent particles, transferring angular momentum by a mechanism first observed by Beth in 1936⁴, when he reported a tiny torque developed in a quartz 'wave-plate' owing to the change in polarization of transmitted light. Here we show that an optical torque can be induced on microscopic birefringent particles of calcite held by optical tweezers⁵. Depending on the polarization of the incident beam, the particles either become aligned with the plane of polarization (and thus can be rotated through specified angles) or spin with constant rotation frequency. Because these microscopic particles are transparent, they can be held in three-dimensional optical traps at very high power without heating, leading to rotation rates of over 350 Hz.

A typical optical-tweezers arrangement was used to trap microscopic calcite particles in three dimensions using between 30 and 300 mW of laser light at a wavelength of 1,064 nm. The optical trap used a 100× oil-immersion, high numerical aperture (NA = 1.3) microscope objective. The trapping beam was initially linearly polarized, and the plane of polarization could be rotated using a half-wave plate. Alternatively, a quarter-wave plate allowed the ellipticity of polarization to be varied. The particles were fragments obtained by crushing a small crystal, giving irregular particles 1–15 μm across. They were dispersed in distilled water in a trapping cell consisting of a well in a microscope slide with a coverslip.

Because of their birefringent nature, calcite particles can act as wave-plates; a calcite particle 3 μm thick is a λ/2 plate for 1,064-nm light. On passage through a fragment of calcite, the ordinary and extraordinary components of the incident light will undergo different phase shifts. If this results in a change in the angular momentum carried by the light, there will be a corresponding torque on the material. Our results can be understood using a simple plane-wave picture; the interaction between an incident plane wave and a wave-plate is outlined below. We note that the calcite wave-plate is trapped at the focal point of the beam, where the wavefronts are nearly plane.

An incident laser beam can in general have both circularly polarized and plane polarized components; that is, it will be elliptically polarized. Elliptically polarized light can be described by $\mathbf{E} = E_0 e^{i\omega t} \cos \phi \hat{x} + i E_0 e^{i\omega t} \sin \phi \hat{y}$ where ϕ describes the degree of ellipticity of the light ($\phi = 0$ or $\pi/2$ indicates plane-polarized light, $\phi = \pi/4$ circularly polarized light). The angular momentum of a plane electromagnetic wave (the incident light) of angular frequency ω can be found from the electric field $\mathbf{E}$ and its complex conjugate $\mathbf{E}^*$ by integrating over all spatial elements d^3r giving $\mathbf{J} = (\epsilon/(2i\omega)) \int d^3r \mathbf{E}^* \times \mathbf{E}$, where ϵ is the permittivity.

To calculate the change in angular momentum of the light after passage through a birefringent material, the incident elliptically polarized light is first expressed in terms of components parallel and perpendicular to the optic axis of the material by:

$$\mathbf{E} = E_0 e^{i\omega t} (\cos \phi \cos \theta - i \sin \phi \sin \theta) \hat{i} + E_0 e^{i\omega t} (\cos \phi \sin \theta + i \sin \phi \cos \theta) \hat{j} \quad (1)$$

where θ is the angle between the fast axis of the quarter-wave plate producing the elliptically polarized light and the optic axis of the birefringent material. The phase shift due to passing through a thickness d with refractive index n is kdn , where k is the free-space wavenumber, so the emergent light field will be

$$\mathbf{E} = E_0 e^{i\omega t} e^{ikdn_e} (\cos \phi \cos \theta - i \sin \phi \sin \theta) \hat{i} + E_0 e^{i\omega t} e^{ikdn_o} (\cos \phi \sin \theta + i \sin \phi \cos \theta) \hat{j} \quad (2)$$

where n_e and n_o are the refractive indices of the birefringent material experienced by extraordinary and ordinary rays, respectively.

The changes in the angular momentum of the light cause a reaction torque per unit area on the thickness d of material of:

$$\tau = -\frac{\epsilon}{2\omega} E_0^2 \sin(kd(n_o - n_e)) \cos 2\phi \sin 2\theta + \frac{\epsilon}{2\omega} E_0^2 \{1 - \cos(kd(n_o - n_e))\} \sin 2\phi \quad (3)$$

In general, the first term of equation (3) is the torque due to the plane-polarized component of elliptically polarized light while the second term is due to the change in polarization caused by passage through the medium. For plane-polarized light, $\phi = 0$ or $\pi/2$, so the torque on the particle is proportional to $\sin 2\theta$, so that a particle will experience torque so long as θ is non-zero, and will be at equilibrium when the fast axis of the crystal is aligned with the plane of polarization ($\theta = 0$). We found that calcite fragments trapped in plane-polarized light are aligned in a particular orientation, and a particular particle is always aligned in the same plane each time it is trapped. When the plane of polarization is rotated using a half-wave

plate, a particle's alignment exactly follows the rotation of the plane of polarization. In Fig. 1, a calcite fragment is shown to rotate through 80° as a half-wave plate controlling the polarization of the trapping beam is rotated through 40° , illustrating the alignment of birefringent particles to the plane of polarization. To our knowledge, this is the first report of an optically trapped particle being rotated through a preset angle; a modification to the set-up whereby the half-wave plate could be spun at a set rate would also allow the particle to rotate at a preset frequency.

The second term of equation (3) will be constant for a given laser power and ellipticity of polarization (characterized by ϕ), and will be maximum for circularly polarized light when the first term vanishes. Hence, when trapped in a circularly polarized beam, a birefringent particle will experience constant torque. In a viscous medium, this torque will be balanced by the drag torque, $\tau_D = D\Omega$, where D is the drag coefficient and Ω is the angular speed, so in this case a birefringent particle will rotate with constant frequency and angular speed. We measured the rotation frequencies of trapped

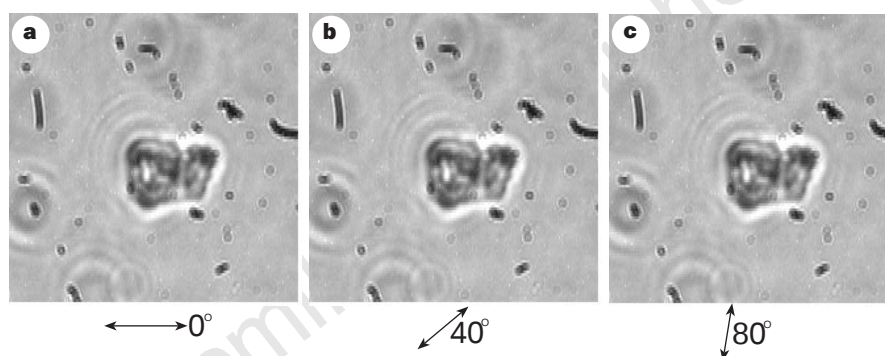


Figure 1 Three sequential photographs (frames) of a trapped calcite crystal, showing alignment with the plane of polarization of the trapping beam. A $\lambda/2$ plate was rotated by 20° between successive photographs, rotating the plane of

polarization by 40° , as shown by the arrows, and exerting an alignment torque on the crystal, causing it to rotate to a new position. This can be used to rotate the particle at a controlled speed, or to control its orientation.

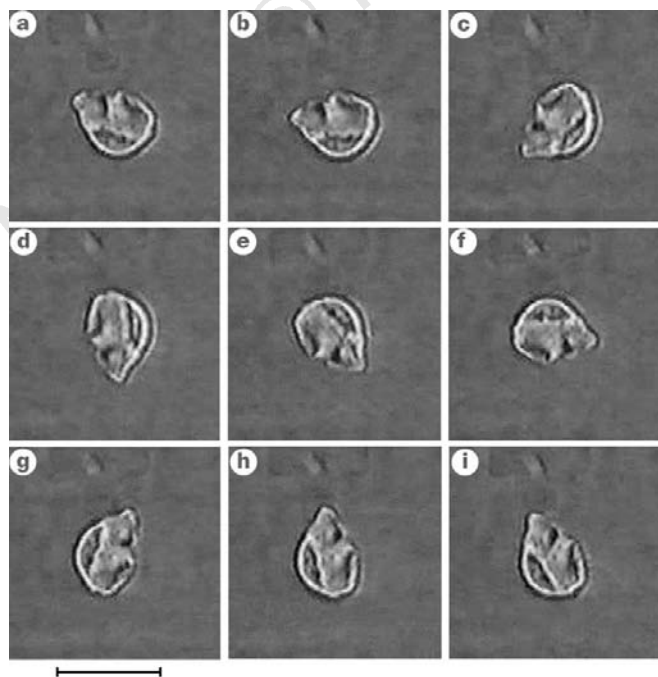


Figure 2 Nine frames of a trapped calcite crystal, showing free rotation due to an elliptically polarized trapping beam. The speed of rotation is limited by the viscous drag on the particle. As the optical torque acting on the particle depends on its orientation, the rotation speed is not constant. The frames are 40 ms apart. Scale bar, 10 μm .

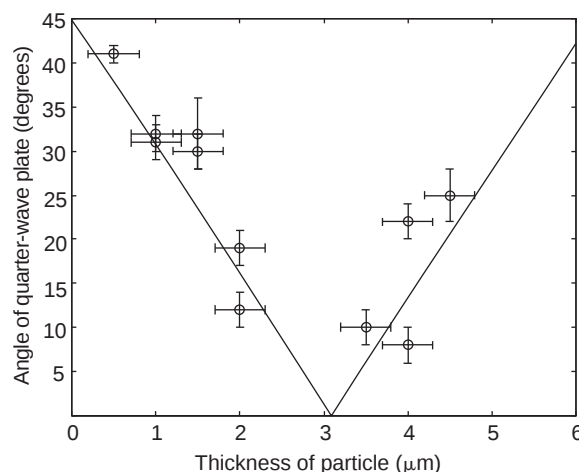


Figure 3 The degree of circular polarization required to cause spinning of a trapped particle depends on its thickness. Here we compare measurements of the minimum angle required for rotation with the theoretical solution. Only if the particle is the exact thickness of a half-wave plate will it always spin in elliptical light. For all other particle thicknesses, there is some angle θ for which the maximum alignment torque will be greater than the spinning torque. The degree of ellipticity of polarization (measured by ϕ) required for the onset of rotation is found from the case where the alignment torque is maximum and the total torque is zero, which is when $\sin[kd(n_o - n_e)] \cos 2\phi = \{1 - \cos[kd(n_o - n_e)]\} \sin 2\phi$. The solution to this is $\phi_{\text{rotate}} = [\pi - kd(n_o - n_e)]/4$. In general a particle will be aligned to the plane of polarization of the trapping beam unless there is sufficient torque due to the circularly polarized component to set it into rotation, when it will experience a position-dependent torque.

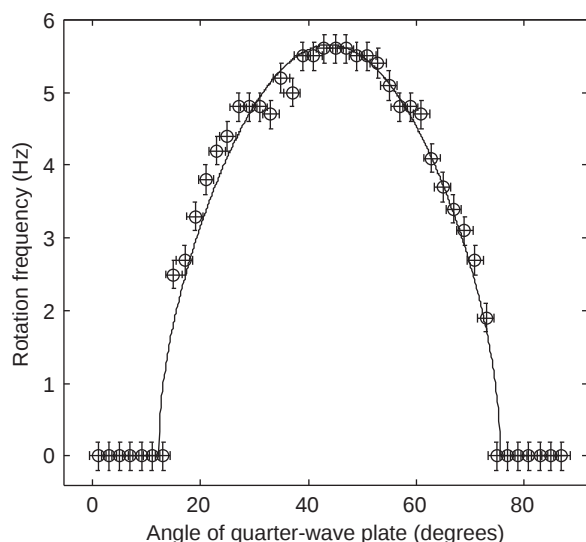


Figure 4 The variation of rotation frequency with the polarization of the trapping beam. The sudden onset of rotation when the minimum angle is reached can be clearly seen. The theoretical response of the particle allows a particle of unknown size to be measured, or can be used to determine the viscosity of a fluid. In this case, the trapping beam has a power of 50 mW, and the particle is 6.15 μm in radius and 2.3 μm thick. The frequency of rotation for a trapping beam of power P and optical frequency ω is $f = [\text{Re}\{P[1 - \cos kd(n_o - n_e)]^2 \sin^2 2\phi - \sin^2 kd(n_o - n_e) \cos^2 2\phi\}^{1/2}]/(2\pi\omega D)$. The drag torque coefficient D can be estimated by representing the particle as an ellipsoid⁶. The drag will lie between that of a sphere of radius a in a medium of viscosity μ ($D = 8\pi\mu a^3$) and a disk of the same radius ($D = (32/3)\mu a^3$). Under the very low Reynolds number flow conditions encountered here, the surface texture and fine structure of the particle are unimportant. The maximum rotation speed, $f = P/\pi\omega D$, will result when the incident light is circularly polarized and the particle is a half-wave plate. Small particles will generally rotate faster due to less drag, but as particles become too small, their thickness becomes much less than the ideal half-wave case, and they will not intercept all of the power available to spin larger particles.

calcite fragments by detection of back-scattered light¹; the results show that calcite fragments rotate at constant frequency in circularly polarized light, and that this frequency is proportional to the laser power. A rotating particle is shown in Fig. 2. The fastest rotation frequency measured was 357 Hz, for a particle 1 μm thick trapped in a 300-mW laser beam.

For the general case of elliptically polarized light, both the alignment torque and the spinning torque will act, and the effect on the birefringent particle will depend on the thickness d and the ellipticity ϕ of the light. The particle will only rotate if the maximum alignment torque is less than the spinning torque. This is shown in Fig. 3. In Fig. 4 we plot the variation of rotation rate of a larger calcite crystal with degree of ellipticity of the trapping beam ϕ , showing the characteristic behaviour of a birefringent particle in elliptically polarized light.

The agreement between our results and the theory outlined above shows that the calcite particles act as microscopic wave-plates. The measurement of the rotation speed of spinning particles is a less accurate but simpler analogue of Beth's experiment⁴. In particular, assuming conventional viscous drag, the observed speeds are consistent with the accepted intrinsic spin angular momentum of $\hbar$ per photon (see Fig. 4). Our results also show how optical torques can be exerted on certain microscopic objects with high precision and efficiency, with minimal heating. Depending on the details of the arrangement, a constant torque independent of orientation can be exerted, leading to rotation rates up to hundreds of hertz, or alternatively the orientation of the object can be smoothly controlled.

The controllability of the motion of calcite particles and the minimal absorption involved suggests calcite as an ideal material for optically driven rotary micromachines. Such micromachines could include pumps, stirrers, or optically powered cogwheels. The rotation could also be used to study the viscosity of small samples of fluids, or the alignment of calcite particles could be used to hold probe particles in particular orientations, which could be useful for atomic-force or other forms of microscopy. The birefringence of biological samples is usually much less than that of calcite, but may sometimes be large enough to allow the same alignment and free rotation to be achieved. □

Received 12 December 1997; accepted 7 May 1998.

1. Friesen, M. E. J., Enger, J., Rubinsztajn-Dunlop, H. & Heckenberg, N. R. Optical angular-momentum transfer to trapped absorbing particles. *Phys. Rev. A* **54**, 1593–1596 (1996).
2. He, H., Friesen, M. E. J., Heckenberg, N. R. & Rubinsztajn-Dunlop, H. Direct observation of transfer of angular momentum to absorptive particles from a laser beam with a phase singularity. *Phys. Rev. Lett.* **75**, 826–829 (1995).
3. Simpson, N. B., Dholakia, K., Allen, L. & Padgett, M. J. Mechanical equivalence of the spin and orbital angular momentum of light: an optical spanner. *Opt. Lett.* **22**, 52–54 (1997).
4. Beth, R. A. Mechanical detection and measurement of the angular momentum of light. *Phys. Rev.* **50**, 115–125 (1936).
5. Ashkin, A., Dziedzic, J. M., Bjorkholm, J. E. & Chu, S. Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett.* **11**, 288–290 (1986).
6. Constantinou, V. N. *Laminar Viscous Flow* (Springer, New York, 1995).

Acknowledgements. This work was supported by the Australian Research Council.

Correspondence and requests for materials should be addressed to T.A.N. (e-mail: timo@physics.uq.edu.au).

High-throughput screening of solid-state catalyst libraries

Selim M. Senkan

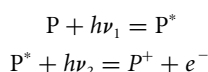
Department of Chemical Engineering, University of California, Los Angeles, California 90095-1592, USA

Combinatorial synthesis methods allow the rapid preparation and processing of large libraries of solid-state materials. The use of these methods, together with the appropriate screening techniques, has recently led to the discovery of materials with promising superconducting¹, magnetoresistive², luminescent^{3–5} and dielectric⁶ properties. Solid-state catalysts, which play an increasingly important role in the chemical and oil industries⁷, represent another class of material amenable to combinatorial synthesis. Yet typically, catalyst discovery still involves inefficient trial-and-error processes^{8–10}, because catalytic activity is inherently difficult to screen. In contrast to superconductivity, magnetoresistivity and dielectric properties, which can be tested by contact probes, or luminescence, which can be observed directly, the assessment of catalytic activity requires the unambiguous detection of a specific product molecule above a small catalyst site on a large library. Screening by *in situ* infrared thermography¹¹ and microprobe sampling mass spectrometry^{12,13} have been suggested, but the first method, while probing activity, provides no information on reaction products, whereas the second is difficult to implement because it requires the transport of minute gas samples from each library site to the detection system. Here I describe the use of laser-induced resonance-enhanced multiphoton ionization for sensitive, selective and high-throughput screening of a library of solid-state catalysts that activate the dehydrogenation of cyclohexane to benzene. I show that benzene, the product molecule, can be selectively photoionized in the vicinity of the catalytic sites, and that the detection of the resultant photoions by an array of microelectrodes provides information on the activity of individual sites. Adaptation of this technique for the screening of other catalytic reactions and larger libraries with smaller site size seems feasible, thus opening up the possibility of exploiting combinatorial

approaches as an efficient method for the discovery and optimization of solid-state catalysts.

The screening method described here exploits tunable lasers for the selective photoionization of product molecules in the vicinity of catalytic sites, followed by the detection of the resulting photoions or photoelectrons by an array of microelectrodes. When the laser frequency is tuned to a real intermediate electronic state of a molecule, the cross-section for ionization is significantly enhanced. This is called resonance-enhanced multiphoton ionization (REMPI; ref. 14). If the laser is not tuned to a real electronic state, the probability for photoionization is smaller. Consequently, a desired reaction product can be selectively ionized with high efficiency while avoiding the simultaneous photoionization of the reactants, other products and background gases provided that the absorption spectrum of the latter molecules is sufficiently different from that of the reaction product of interest.

The most common REMPI approach is resonant two-photon ionization, in which one photon ($h\nu_1$) energizes the molecule P to an excited electronic state and the second photon ($h\nu_2$) ionizes the molecule:

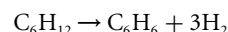


where P^* is a real electronic excited state, P^+ is the photoion and e^- is the photoelectron. The absorption of two or more photons in each step may also be utilized but this may require the laser beam to be focused. Ionization occurs if $(h\nu_1 + h\nu_2) > \text{IP}$, where IP is the ionization potential. By varying the photon energies, the ionization of P (as well as that of other molecules of interest) can be mapped out to determine suitable laser frequencies for its exclusive ionization. The two photons used can have the same or different frequencies. The REMPI process can also be repeated to sequentially detect different products (using different laser frequencies) for the

determination of product selectivities. REMPI is a sensitive technique with real-time detection at low parts per billion (ref. 15) and high parts per trillion (ref. 16) already demonstrated. However, each application of REMPI requires a careful analysis of the spectral properties of the molecules of interest. If these properties are not known, they will need to be determined in separate experiments. It will be difficult to selectively identify a molecule in the presence of another species that absorbs in the same wavelength region. Molecules that dissociate to form excited states and/or have high IPs (such as methanol or ethanol) will also be difficult to probe; that is, a signal can only be obtained by focusing the laser to induce a multiphoton transition that avoids the lower dissociative states. However, many aromatic systems, aldehydes and ketones can be probed with high sensitivity and selectivity using REMPI.

In Fig. 1 I show a sketch of the overall system that has been developed to screen catalyst libraries. Details of the reactant and catalyst contact, and the relative positions of the laser beam, the microelectrodes and the catalyst sites are shown in Fig. 2. REMPI signals are detected by the use of dedicated microelectrodes placed above each site.

The technical feasibility of the screening method was demonstrated by considering the catalytic dehydrogenation of cyclohexane into benzene:



This is an established reaction that is catalysed by transition and precious metals in the temperature range 250–350 °C (refs 17, 18). Pt and Rh catalysts were acquired from a commercial vendor (Precious Metals Corp., TN). Catalysts were crushed into 0.2–0.5 mm pieces, and batches of ~40 mg were manually placed in the fifth row of the library, as shown in Figs 1 and 2. The catalyst library was then placed into the reactor and the entire system was heated to 300 °C in the presence of argon gas flow. The feed stream

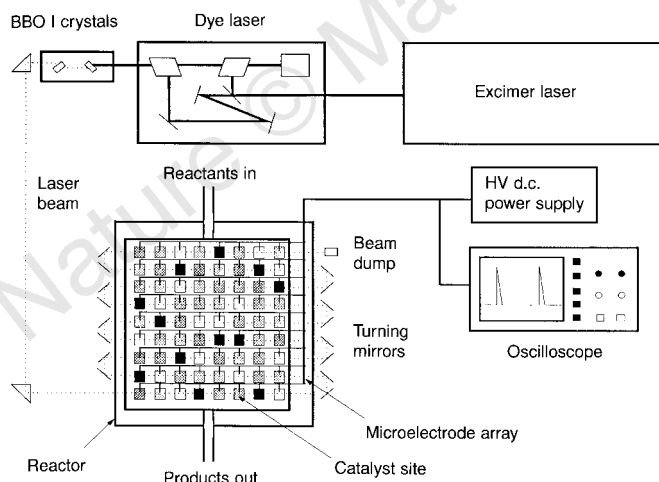


Figure 1 Experimental set-up for the screening of solid-state catalyst libraries. The system consists of a reactor chamber, a temperature-controlled furnace and the tunable ultraviolet laser source. The library consisted of a square ceramic substrate (7.5 × 7.5 cm) containing 72 sites (8 by 9 rows). Each site was 5 × 5 mm and separated 2 mm from one another, corresponding to a library density of ~10 sites per square inch. The reactor was designed to introduce the reactant gases gently, and to remove the products from the vessel. Photoionization was achieved by passing a pulsed, tunable ultraviolet laser beam over all the sites in a row, as indicated. Suitable laser light frequencies were obtained from a dye laser (LDL 2051, Laser Analytical Systems, Santa Clara, CA) that was pumped by an excimer laser (Complex 102, Lambda Physik, Ft Lauderdale, FL). A BBO I crystal was used for second-harmonic generation (BBO, beta barium borate). Optical access to the library was provided through MgF₂ windows placed at the side walls of the reactor.

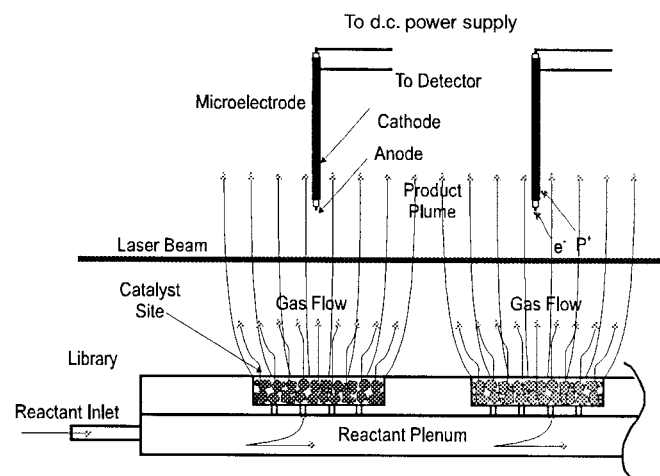


Figure 2 Design of individual sites on the catalyst library. The reactant gases are forced through individual sites, generating product plumes if the sites are catalytic. The laser beam is then passed through the air-spaces above all the sites, intercepting the plumes, and thereby generating the product photoions and photoelectrons. Dedicated microelectrodes placed above the sites in the close vicinity of the laser beam simultaneously collect either the photoions or photoelectrons, depending on the d.c. bias voltage applied. The use of other library geometries, such as those based on monolithic or honeycomb structures with well defined channels, are also feasible. Microelectrodes were 0.5 mm in diameter, 3 cm long, and possessed both the anode and the cathode¹⁹. They were mounted through a ceramic plate that was attached to the top of the reactor chamber. A multichannel switch was used to individually apply d.c. power (PS 350, Stanford Research Systems, Sunnyvale, CA) to each microelectrode and to screen the sites. Signal detection was accomplished by a digital oscilloscope (Hewlett-Packard 54540C).

composition was then changed to 13% cyclohexane in argon carrier gas, prepared by bubbling argon through cyclohexane at $\sim 25^\circ\text{C}$.

The screening protocol demands the unambiguous detection of benzene in cyclohexane, hydrogen and argon background. In order to accomplish this goal one must identify a suitable ultraviolet laser wavelength that selectively produces benzene REMPI ions. This wavelength was identified using a time-of-flight mass spectrometer (TOF-MS). In Fig. 3a the TOF-MS photoionization spectra of benzene and cyclohexane are presented: benzene shows large REMPI peaks starting at 259.7 nm, and there are no contributions from cyclohexane. Following the TOF-MS experiments, REMPI spectra were also determined at ambient conditions using the microelectrode detection system. As seen in Fig. 3b, the REMPI spectra obtained by the microelectrode technique were similar to the TOF-MS spectra, with the expected broadening in room-temperature experiments. Figure 3b also shows that a significant

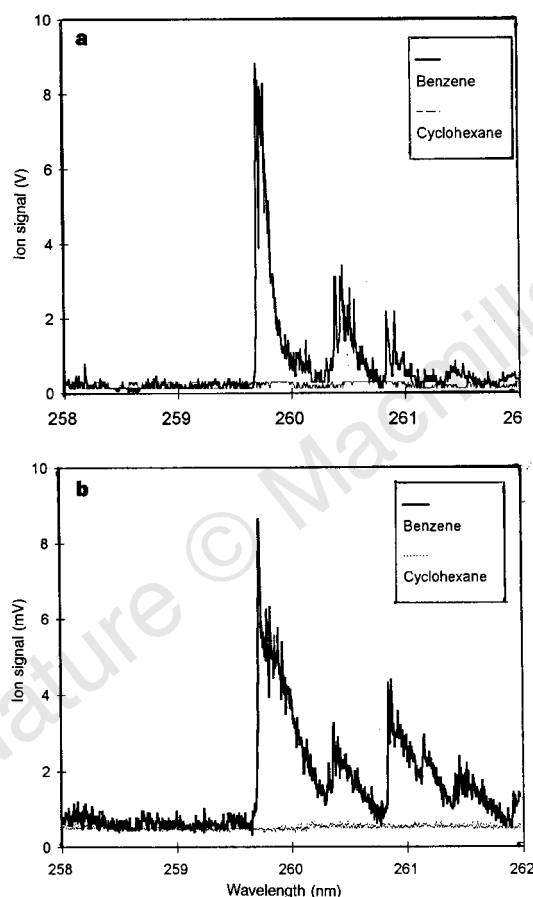


Figure 3 Control spectra for the cyclohexane/benzene system. **a**, Cyclohexane and benzene, each at a concentration of ~ 500 p.p.m. in argon, were introduced into the vacuum chamber of the TOF-MS using a pulsed valve, and the resulting jet/molecular beam was crossed by a pulsed ultraviolet laser light in the 258–262 nm range to generate their photoionization spectra. The ultraviolet light had an energy of ~ 100 μJ per pulse over the entire wavelength range, and was obtained by frequency-doubling the beam from the dye laser (see Fig. 1). These measurements led to the conclusion that the REMPI ions produced by 258–262 nm light were exclusively due to the photoionization of benzene (mass 78), with no photoions detected as masses 84 (cyclohexane), 40 (argon) or 2 (hydrogen). Furthermore, no peaks other than the benzene parent as mass 78 were detected. **b**, REMPI spectra obtained under ambient conditions using microelectrode detection. Experiments were done by flowing cyclohexane- or benzene-bearing argon gas over the microelectrode, and photoionizing the gases within 1–2 mm of the probe tip using the same laser source as in the TOF-MS experiments. A d.c. bias of +500 V was applied to the anode to collect the photoelectrons.

broadening of the REMPI spectra can be tolerated in catalyst screening when the REMPI features of the reactants and products are well separated.

Figure 4 shows the results of the catalyst library screening process (using 259.7-nm light) for the fifth row. These measurements correspond to data acquired by one laser shot and do not involve signal averaging. The signals show decay times on the order of microseconds, suggesting the feasibility of screening large libraries. Microelectrodes located at sites 2, 4, 7 and 8 picked up appreciable benzene REMPI signals, consistent with the presence of Pt and Pd catalysts at these sites.

I note that although some benzene REMPI signals were detected at sites 1, 3, 5 and 6, they were significantly lower, consistent with the absence of catalysts at these sites. Evidently some benzene remained in the reactor because of the low gas-flow rates used and the recirculation patterns present, both of which prevent the rapid removal of products from the reactor. A smaller chamber design and higher gas-flow rates, or the use of monolithic or honeycomb channel structures, could minimize this problem. Nevertheless, the technique clearly distinguished between the active and inactive sites on the catalyst library. The reactor exhaust gases were also analysed by the TOF-MS using the 259.7-nm laser light to find out if species other than benzene could have contributed to the measured signals. No photoions, other than those with mass 78, were detected. Photoionization MS analysis of the reactor exhaust gases by the same laser used in the screening process ensures that microelectrode measurements are exclusively due to the desired product and are not due to the formation of another by-product catalysed by the library.

Based on the magnitudes of the REMPI signals measured (Fig. 4), the relative activities of the catalytic sites appear to be $7 > 2 > 4 > 8$. Although fortuitous, as the microelectrodes were not calibrated, these results are consistent with the relative loadings of the Pd and Pt in catalysts used. In addition, they also suggest that Pt is a more active dehydrogenation catalyst than Pd, in good agreement with the results of earlier studies using conventional catalytic reactor systems^{17,18}.

These results indicate that quantitative data on the activities of each site in the catalyst library can be obtained, provided each microelectrode system is calibrated. This calibration can be done electronically to produce the same signal intensity when using a calibration gas mixture (or mixtures) in the absence of the catalyst, that is, no reaction. During the actual library screening process, the signals measured will be proportional to the activities of the catalyst sites, but also include background contributions. The latter can be corrected for by using signals measured at known non-catalytic

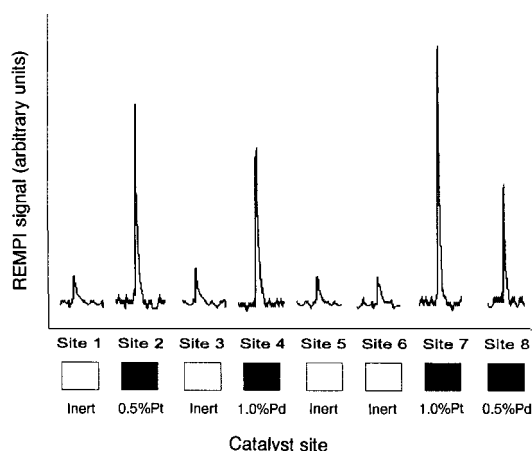


Figure 4 Screening results of the fifth row in the catalyst library. The addresses of the various catalysts used in the experiments were: site 1, inert; site 2, 0.5% Pt; site 3, inert; site 4, 1.0% Pd; site 5, inert; site 6, inert; site 7, 1.0% Pt; site 8, 0.5% Pd.

sites. Clearly, differences in catalyst loading must also be taken into account when assessing the relative activities of the catalytic sites.

The technique described here could be extended to monitor multiple reaction products, and thus could also be used to acquire information on catalyst selectivity. This could be accomplished by using different laser frequencies to sequentially generate the REMPI signals of different products. The REMPI signals could then be converted into absolute concentrations, using calibration standards, for the determination of selectivities. In addition, the technique developed should be useful in the study of issues related to the operational lifetimes of catalysts, their resistance to poisoning, their regeneration and their loss during operation. □

Received 26 March; accepted 5 May 1998.

- Xiang, X.-D. *et al.* A combinatorial approach to materials discovery. *Science* **2268**, 1738–1740 (1995).
- Briceno, G., Chang, H., Sun, X., Schultz, P. G. & Xiang, X.-D. A class of cobalt oxide magnetoresistance materials discovered with combinatorial synthesis. *Science* **270**, 273–275 (1995).
- Sun, X.-D. *et al.* Solution-phase synthesis of luminescent materials libraries. *Adv. Mater.* **9**, 1046–1049 (1997).
- Sun, X.-D., Gao, C., Wang, J. & Xiang, X.-D. Identification and optimization of advanced phosphors using combinatorial libraries. *Appl. Phys. Lett.* **70**, 3353–3355 (1997).
- Danielson, E. *et al.* A combinatorial approach to the discovery and optimization of luminescent materials. *Nature* **389**, 944–948 (1997).
- van Dover, R. B., Schneemeyer, L. F. & Fleming, R. M. Discovery of a useful thin-film dielectric using a composition-spread approach. *Nature* **392**, 162–164 (1998).
- National Research Council *Catalysis Looks to the Future* (National Academy Press, Washington DC, 1992).
- Poncelet, G., Martens, J., Delmon, B., Jacobs, P. A. & Grange, P. *Preparation of Catalysts VI* (Elsevier, Amsterdam, 1995).
- Heinemann, H. in *Catalysis: Science and Technology* (eds Anderson, J. R. & Boudart, M.) Ch. 1 (Springer, Berlin, 1981).
- Satterfield, C. N. *Heterogeneous Catalysis in Practice* 2nd edn (McGraw Hill, New York, 1991).
- Moates, F. C. *et al.* Infrared thermographic screening of combinatorial libraries of heterogeneous catalysts. *Ind. Eng. Chem. Res.* **35**, 4801–4803 (1996).
- Kassem, M. & Senkan, S. M. Chemical structure of fuel-rich 1,2-C₂H₄Cl₂/CH₄/O₂/Ar flames: effects of microprobe cooling on the sampling of chlorinated hydrocarbons. *Combust. Sci. Technol.* **67**, 147–157 (1989).
- Service, R. F. High-speed materials design. *Science* **277**, 474–475 (1997).
- Lubman, D. & Li, L. in *Lasers and Mass Spectroscopy* (ed. Lubman, D. M.) Ch. 16 (Oxford Univ. Press, New York, 1990).
- Gittins, C. M., Castaldi, M. J., Senkan, S. M. & Rohlfing, E. A. Real-time quantitative analysis of combustion generated polycyclic aromatic hydrocarbons by resonance enhanced multiphoton ionization time of flight mass spectrometry. *Anal. Chem.* **69**, 286–293 (1997).
- Senkan, S. M. & Castaldi, M. J. Real-time, ultrasensitive monitoring of air toxics by laser photoionization time of flight mass spectrometry. *J. Air Waste Mgmt. Assoc.* **48**, 77–81 (1998).
- Rehbon, D. M. & Haensel, V. A kinetic and mechanistic study of cyclohexane disproportionation: an example of irreversible hydrogen transfer. *J. Catal.* **111**, 397–408 (1988).
- Ahmed, K. & Chowdhury, H. M. Dehydrogenation of cyclohexane and cyclohexene over supported nickel and platinum catalysts. *Chem. Eng. J.* **50**, 165–168 (1992).
- Cool, T. A. Quantitative measurement of NO density by resonance three-photon ionization. *Appl. Opt.* **23**, 1559–1572 (1984).

Acknowledgements. I thank LCS Inc. for funding and use of their facilities.

Correspondence and requests for materials should be addressed to the author (e-mail: senkan@seas.ucla.edu).

Unexpectedly high concentrations of molecular chlorine in coastal air

C. W. Spicer*, E. G. Chapman†, B. J. Finlayson-Pitts‡, R. A. Plastringer*, J. M. Hubbe†, J. D. Fast† & C. M. Berkowitz†

* Battelle, 505 King Avenue, Columbus, Ohio 43201-2693, USA

† Pacific Northwest National Laboratory, PO Box 999, Richland, Washington 99352-2122, USA

‡ University of California, Irvine, California 92697-2025, USA

The fate of many atmospheric trace species, including pollutants such as nitrogen oxides and some volatile organic compounds, is controlled by oxidation reactions. In the daytime troposphere, these reactions are dominated by photochemically produced OH radicals; at night and in polluted environments, NO₃ radicals are an important oxidant¹. Ozone can contribute to the oxidation of atmospheric species during both day and night¹. In recent years,

laboratory investigations^{2–4}, modelling studies^{5–7}, measured Cl deficits in marine aerosols⁸ and species-nonspecific observations^{9–11} of gaseous inorganic chlorine compounds other than HCl have suggested that reactive halogen species may contribute significantly to—or even locally dominate—the oxidative capacity of the lower marine troposphere. Here we report night-time observations of molecular chlorine concentrations at a North American coastal site during onshore wind flow conditions that cannot be explained using known chlorine chemistry. The measured Cl₂ mixing ratios range from <10 to 150 parts per 10¹² (p.p.t.), exceeding those predicted⁵ for marine air by more than an order of magnitude. Using the observed chlorine concentrations and a simple photochemical box model, we estimate that a hitherto unrecognized chlorine source must exist that produces up to 330 p.p.t. Cl₂ per day. The model also indicates that early-morning photolysis of molecular chlorine can yield sufficiently high concentrations of chlorine atoms to render the oxidation of common gaseous compounds by this species 100 times faster than the analogous oxidation reactions involving the OH radical, thus emphasizing the locally significant effect of chlorine atoms on the concentrations and lifetimes of atmospheric trace species in both the remote marine boundary layer and coastal urban areas.

To determine if reactive halogens exist in the marine boundary layer (MBL), an atmospheric-pressure chemical ionization tandem mass spectrometer (Sciex Model 6000E trace atmospheric gas analyser) was used to measure Cl₂. The instrument provides high selectivity by using two stages of mass analysis: the parent ion—an intact molecular ion selected in the first mass analyser—undergoes collisionally activated dissociation to yield fragments (daughter ions) whose mass-to-charge ratios (*m/z*) can be determined by the second mass analyser. Highly selective monitoring then can be performed by using the two mass filters to pass only the parent and daughter ions corresponding to the species of interest. Chlorine has two natural isotopes with masses 35 and 37 a.m.u. and Cl₂ therefore gives rise to four different parent/daughter ion pairs which were continuously monitored at *m/z* ratios of 70/35, 72/35, 72/37 and 74/37. Calibrations were performed using certified Cl₂ permeation sources whose emission rates were quantified using ion chromato-

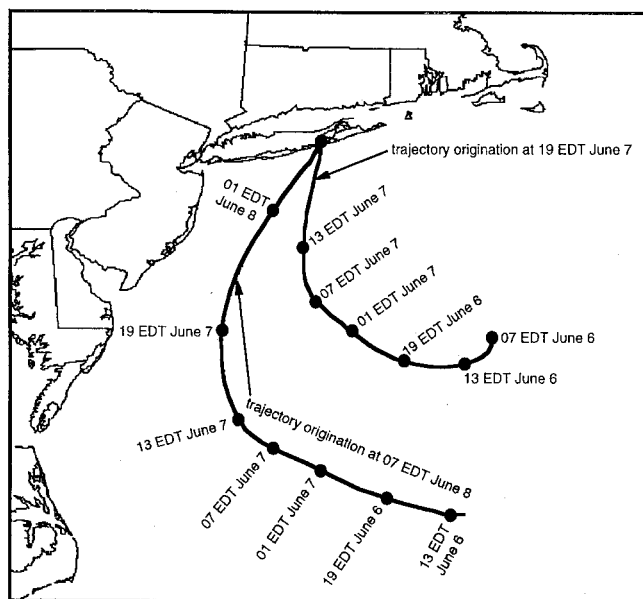


Figure 1 Calculated back trajectories. These are based on mesoscale meteorological wind fields originating at 19:00 EDT on 7 June 1996 (right trajectory) and 07:00 EDT on 8 June 1996 (left trajectory) for air reaching the coastal monitoring site. Circles denote air mass positions at 6-h intervals. All back trajectory positions during the evening lie between these two trajectories.

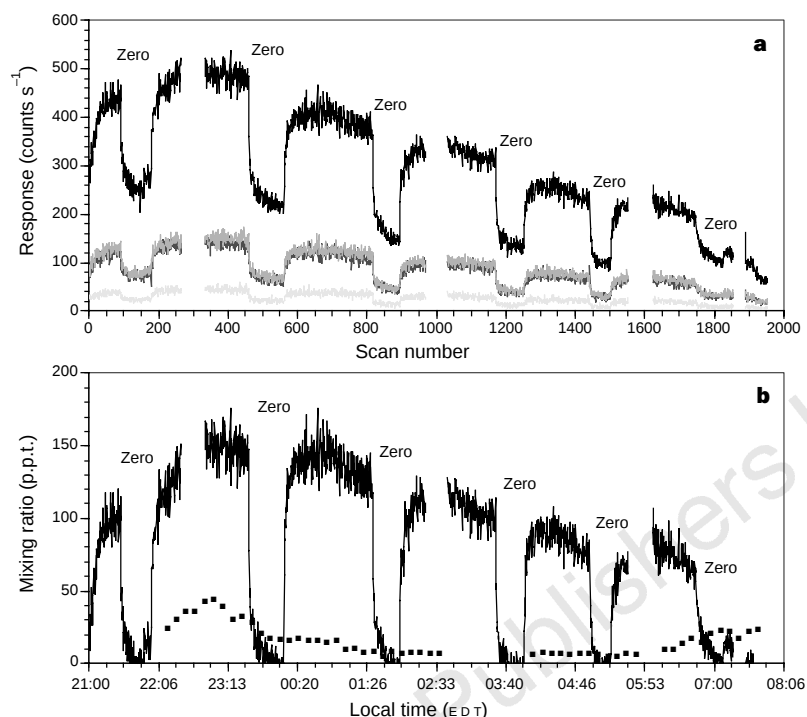


Figure 2 Cl_2 observations at the coastal Long Island, New York site ($40^\circ 51' \text{N}$, $72^\circ 37' \text{W}$) on the evening of 7–8 June 1996. **a**, The raw signal response in counts per second for monitored ion pairs; from top to bottom, 70/35, 72/37, 72/35 (underlies ion pair 72/37) and 74/37. **b**, The corresponding time-averaged Cl_2

mixing ratios (trace) and concurrent two-minute integrated PAN measurements (filled squares). Cl_2 calibration intervals are removed, and periods of instrument zero checks are noted.

graphy, while the purity of the emitted Cl_2 was assessed using mass spectrometry. Instrument background signal was determined using an annular denuder coated with Na_2CO_3 to remove molecular halogens from the air stream while passing other gaseous species and particles to the mass spectrometer inlet. The Cl_2 detection limit is typically 15 p.p.t. Precision is estimated at 10% at mixing ratios near the detection limit, and the accuracy of the calibration method at 20%. Sample air is transported to the mass spectrometer inlet at a rate of 8 l min^{-1} through 2 m of Teflon tubing (6.3 mm outer diameter). A valve directs the air to either pass through or bypass the annular denuder. The atmospheric-pressure ionization chamber of the mass spectrometer is heated to 120°C . The data have been corrected for experimentally determined Cl_2 loss in the sampling system (30%). Further details of the instrument, its modifications for atmospheric research, and its use for monitoring other atmospheric species are given in refs 12–14.

Cl_2 measurements were made in early June 1996 at a ground site in eastern Long Island, New York. The site was located 2.8 km north of the Atlantic coast. The most extensive measurements were made from the evening of 7 June 1996 until the end of the next morning. Surface winds in the vicinity of Long Island and at offshore buoys were from the south during the evening of 7 June, becoming

southwesterly during the morning of 8 June. Fog was observed at the site during portions of the sampling period. Back trajectories (Fig. 1) clearly show that the sampled air mass was of marine origin, having traversed open ocean for at least 36 h before reaching the Long Island coast. Forward trajectory analyses conducted using a combined lagrangian particle dispersion-mesoscale meteorological model¹⁵ indicate that the air from the surface up to 900 m above ground level was not influenced by continental emission sources. Onshore flow of marine air, minimally modified by passing over 2.8 km of rural Long Island, thus existed for the entire sampling period. Concurrent measurements of peroxyacetyl nitrate (PAN) and ozone were respectively <50 p.p.t. and <30 parts per billion (p.p.b.), consistent with values typical for marine air.

Figure 2a shows unprocessed signals of the four Cl_2 parent/daughter ion pairs, which unambiguously confirm the presence of Cl_2 . The ion ratios are consistent with the natural isotopic abundances of chlorine. As shown in Fig. 2b, night-time Cl_2 mixing ratios ranged from ~ 40 to ~ 150 p.p.t. during the period 21:00 on 7 June to 07:00 on 8 June, dropping rapidly after sunrise to <15 p.p.t. in full daylight. (Times are given as Eastern Daylight Time, EDT.) Short-term monitoring on other nights with similar onshore airflow showed Cl_2 levels in the range of 16 to 50 p.p.t., as noted in Table 1. These night-time mixing ratios are similar to the species-nonspecific mist-chamber observations at a Florida coastal site of <26 to 254 p.p.t. Cl (corresponding to <13 to 127 p.p.t. if in the form of Cl_2)⁹, and to Arctic measurements of total photolysable chlorine of <9 to 100 p.p.t. as Cl_2 (ref. 10). Our measurements significantly exceed recent model-predicted steady-state concentrations of 0.4–1.5 p.p.t. Cl_2 under marine conditions⁵.

Heterogeneous NaCl reactions are known to produce photochemically active compounds that photolyse to release Cl atoms^{2–4}. A gas-phase photochemical box model, modified to include such heterogeneous reactions, was used to determine the magnitude of the source needed to produce the observed Cl_2 mixing ratios. The

Table 1 Chlorine measurements

Date (in 1996)	Time (EDT)	Cl_2 (p.p.t.)	O_3 (p.p.b.)	PAN (p.p.t.)	Comments
7 June	01:00–02:00	35–40	58	—	Clear
7–8 June	21:00–07:40	<10 –150	18–27	4–43	Occasional fog
8 June	18:23–18:33	16	—	10	Overcast, fog
9 June	17:45–00:00	14–49	19–25	40–46	Overcast
12 June	00:30–01:30	43	21	24	Heavy fog

Ozone and PAN results are shown where available, and generally support the marine origin of the sampled air.

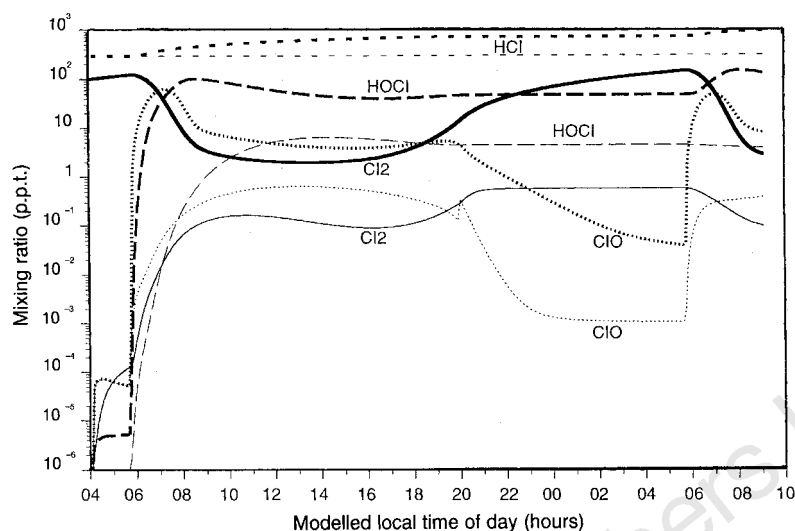


Figure 3 Predicted mixing ratios for chlorine species in air with (bold curves) and without (thin curves) an external source of 330 p.p.t. Cl_2 per day. Initial conditions are typical for marine conditions^{6,21} and include 70 p.p.t. SO_2 , 5 p.p.t. NO_x , 40 p.p.b. O_3 , 20 p.p.t. PAN, 500 p.p.t. C_2H_6 , 15 p.p.t. isoprene, 300 p.p.t. HCl, 110 p.p.t. DMS, 600 p.p.t. CH_3Cl , 150 p.p.t. CH_3CCl_3 , 100 p.p.t. CHClF_2 , 20 p.p.t. C_2Cl_4 , 30 p.p.t. CH_2Cl_2 , 20 p.p.t. CHCl_3 , 20 p.p.t. COCl_2 , and 30 sea-salt particles per cm^3 with diameter of 1 μm . Temperature was 290 K, pressure was 1,014 mbar, and relative

humidity was 99%, consistent with site measurements. Emission fluxes (molecules per cm^2 per s) include ozone 5×10^{10} ; NO 2.5×10^8 ; DMS 2.0×10^9 ; and isoprene, 8.8×10^7 ; these are consistent with marine conditions^{6,21}. Simulations were initiated at 04:00 EDT and run for 29 h, thus including two sunrises. Photolysis rates were specific to the sampling location and date, and were up-dated every 5 min. An average boundary-layer depth of 0.5 km was assumed.

box model takes into account 145 atmospheric species and 292 chemical reactions, including a basic sequence involving hydrocarbons, NO_x and ozone¹⁶, dimethylsulphide (DMS) chemistry^{17,18}, isoprene chemistry^{19–21}, known tropospheric gas-phase chlorine reactions, and the heterogeneous reactions of N_2O_5 , NO_2 , ClONO_2 and HNO_3 with NaCl. The $\text{ClONO}_2 + \text{NaCl}$ reaction is the main source of Cl_2 in this model. All reaction rate coefficients were updated to recent recommendations^{22,23}. Initial conditions and fluxes were based on values typical for a marine air mass^{6,21}.

Model results show that the observed Cl_2 levels cannot be explained by known reactions between NO_x and sea-salt or by gas-phase reactions of the chief Cl-containing species (HCl, CH_3Cl , CH_3CCl_3 , CHClF_2 , C_2Cl_4 , CH_2Cl_2 , CHCl_3 and COCl_2)^{2,3}. A local chlorine source producing an average of 330 p.p.t. Cl_2 per day is required to attain the observed 150 p.p.t. night-time maximum. Figure 3 illustrates the greater build-up of Cl-containing compounds predicted by the model using the source and the effect of excluding it. Maximum predicted Cl atom levels occur shortly after sunrise, peaking at $1.3 \times 10^5 \text{ atoms cm}^{-3}$. Chlorine reacts more rapidly than OH with many trace atmospheric species^{22,23}. Although the modelled OH mixing ratio shortly after sunrise ($3.9 \times 10^5 \text{ molecules cm}^{-3}$) is higher than Cl, the oxidation of many trace species by chlorine still proceeds at a substantially faster rate than by OH. For example, at peak Cl concentrations, the common tropospheric constituent ethane is oxidized by chlorine atoms at a rate of $7.2 \times 10^{-6} \text{ s}^{-1}$, whereas oxidation by the OH radical (whose concentration exceeds that of Cl) proceeds at a rate that is about two orders of magnitude slower ($8.6 \times 10^{-8} \text{ s}^{-1}$). In the case of dimethylsulphide, the rate of oxidation within the model by atomic chlorine shortly after sunrise is an order of magnitude faster than the rate of the corresponding oxidation reaction involving the OH radical.

The contribution of Cl to the oxidizing capacity of the atmosphere is clearly dependent on factors such as time of day, the vertical distribution of oxidants, and the overall chemical composition of an air mass. However, even a lower night-time Cl_2 maximum of 40 p.p.t., requiring a reduced unknown source strength of 88 p.p.t.

Cl_2 per day, leads to post-sunrise ethane oxidation where the reaction with Cl is 25 times faster than the reaction with the OH radical. These simple comparisons suggest, in agreement with Keene *et al.*^{3,8}, the potential importance of chlorine in the atmospheric chemistry of the MBL and coastal urban areas.

The species-specific measurements reported here unambiguously document the presence of substantial amounts of Cl_2 in coastal air masses. Possible sources of molecular chlorine in the MBL include the photolysis of ozone by ultraviolet light, which has been shown to generate Cl_2 at 254 nm and may occur at tropospheric wavelengths of $>290 \text{ nm}$ (ref. 24). Photolytically driven aqueous-phase mechanisms producing gaseous halogens have been proposed^{8,25,26}. However, the present field measurements document the build-up of Cl_2 at night, thus suggesting that another (non-photolytic) mechanism for halogen production may also exist. Additional studies, and associated field measurements, are needed to identify all processes leading to the production of molecular chlorine and further elucidate the overall influence of Cl_2 on tropospheric chemistry. □

Received 20 November 1997; accepted 27 April 1998.

1. Finlayson-Pitts, B. J. & Pitts, J. N. Jr *Atmospheric Chemistry: Fundamentals and Experimental Techniques* (Wiley, New York, 1986).
2. Finlayson-Pitts, B. J. Chlorine atoms as a potential tropospheric oxidant in the marine boundary layer. *J. Res. Chem. Intermed.* **19**, 235–249 (1993).
3. Keene, W. C. in *Naturally-Produced Organohalogens* (eds Grimvall, A. & de Leer, E. W. B.) 363–373 (Kluwer Academic, Norwell, MA, 1995).
4. Behnke, W., Scheer, V. & Zetzsch, C. in *Naturally-Produced Organohalogens* (eds Grimvall, A. & de Leer, E. W. B.) 375–384 (Kluwer Academic, Norwell, MA, 1995).
5. Vogt, R., Crutzen, P. J. & Sander, R. A mechanism for halogen release from sea-salt aerosol in the remote marine boundary layer. *Nature* **383**, 327–330 (1996).
6. Sander, R. & Crutzen, P. J. Model study indicating halogen activation and ozone destruction in polluted air masses transported to the sea. *J. Geophys. Res.* **101**, 9121–9138 (1996).
7. Singh, H. B. & Kasting, J. F. Chlorine-hydrocarbon photochemistry in the marine troposphere and lower stratosphere. *J. Atmos. Chem.* **7**, 261–285 (1988).
8. Keene, W. C. *et al.* The geochemical cycling of reactive chlorine through the marine troposphere. *Glob. Biogeochem. Cycles* **4**, 407–430 (1990).
9. Pszenny, A. A. P. *et al.* Evidence of inorganic chlorine gases other than hydrogen chloride in marine surface air. *Geophys. Res. Lett.* **20**, 699–702 (1993).
10. Impey, G. A., Shepson, P. B., Hastie, D. R., Barrie, L. A. & Anlauf, K. Measurements of photolabile chlorine and bromine during the Polar Sunrise Experiment 1995. *J. Geophys. Res.* **102**, 16005–16010 (1997).
11. Maben, J. R., Keene, W. C., Pszenny, A. A. P. & Galloway, J. N. Volatile inorganic Cl in surface air over eastern North America. *Geophys. Res. Lett.* **22**, 3513–3516 (1995).

12. Spicer, C. W., Kenny, D. V., Shaw, W. J., Busness, K. M. & Chapman, E. G. A laboratory in the sky: New frontiers in measurements aloft. *Environ. Sci. Technol.* **28**, 412A–420A (1994).
13. Chapman, E. G., Kenny, D. V., Busness, K. M., Thorp, J. M. & Spicer, C. W. Continuous airborne measurements of gaseous formic and acetic acids over the western North Atlantic. *Geophys. Res. Lett.* **22**, 405–408 (1995).
14. Spicer, C. W., Kenny, D. V., Chapman, E. G., Busness, K. M. & Berkowitz, C. M. Observations of dimethyl sulfide over the western North Atlantic Ocean using an airborne tandem mass spectrometer. *J. Geophys. Res.* **101**, 29137–29147 (1996).
15. Fast, J. D. & Berkowitz, C. M. A modeling study of boundary-layer processes associated with ozone layers observed during the 1993 North Atlantic Regional Experiment. *J. Geophys. Res.* **101**, 28683–28699 (1996).
16. Lurmann, F. W., Lloyd, A. C. & Atkinson, R. A chemical mechanism for use in long-range transport/acid deposition computer modeling. *J. Geophys. Res.* **91**, 10905–10936 (1986).
17. Yin, F., Grosjean, D. & Seinfeld, J. H. Photooxidation of dimethyl sulfide and dimethyl disulfide. I: Mechanism development. *J. Atmos. Chem.* **11**, 309–364 (1990).
18. Zaveri, R. A. *Development and Evaluation of a Comprehensive Tropospheric Chemistry Model for Regional and Global Applications*. Thesis, Virginia Polytechnic Inst. (1997).
19. Gery, M. W., Whitten, G. Z., Killus, J. P. & Dodge, M. C. A photochemical kinetics mechanism for urban and regional scale computer modeling. *J. Geophys. Res.* **94**, 12925–12956 (1989).
20. Ragains, M. L. & Finlayson-Pitts, B. J. Kinetics and mechanism of the reaction of Cl atoms with 2-methyl-1,3-butadiene (isoprene) at 298K. *J. Phys. Chem.* **101**, 1509–1517 (1997).
21. Bonsang, B., Polle, C. & Lambert, G. Evidence for marine production of isoprene. *Geophys. Res. Lett.* **19**, 1129–1132 (1992).
22. DeMore, W. B. *et al.* *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling* (Rep. 97-4, Jet Propulsion Lab., Pasadena, CA, 1997).
23. Atkinson, R. Gas phase tropospheric chemistry of volatile organic compounds: 1. Alkanes and alkenes. *J. Phys. Chem. Ref. Data* **26**, 215–290 (1997).
24. Oum, K. W., Lakin, M. J., DeHaan, D. O., Brauers, T. & Finlayson-Pitts, B. J. Formation of molecular chlorine from photolysis of ozone and aqueous sea salt particles. *Science* **279**, 74–77 (1998).
25. Chameides, W. L. & Stelson, A. W. Aqueous-phase chemical processes in deliquescent sea-salt aerosols: A mechanism that couples the atmospheric cycles of S and sea salt. *J. Geophys. Res.* **97**, 20565–20580 (1992).
26. Keene, W. C. *et al.* Comment on “Aqueous-phase chemical processes in deliquescent sea-salt aerosols: A mechanism that couples the atmospheric cycles of S and sea salt” by W. L. Chameides and A. W. Stelson. *J. Geophys. Res.* **98**, 9047–9049 (1993).

Acknowledgements. We thank pilots R. V. Hannigan and M. J. Warren for their invaluable help during the field studies; X. Bian for graphics preparation; and W. C. Keene, R. D. Saylor, J. A. Shorter and W. R. Barchet for discussions. We acknowledge the financial support of the US Department of Energy's Atmospheric Chemistry Program.

Correspondence and requests for materials should be addressed to C.W.S. (e-mail: spicer@battelle.org).

Escape tectonics in the Los Angeles metropolitan region and implications for seismic risk

Christian Walls^{*†‡}, Thomas Rockwell^{*†}, Karl Mueller^{§†},
Yehuda Bock^{||†}, Simon Williams^{||†}, John Pfanner^{*†},
James Dolan[†] & Peng Fang^{||†}

^{*} Department of Geological Sciences, San Diego State University, California 92182, USA

[†] Southern California Earthquake Center, University of Southern California, Los Angeles, California 90089-0740, USA

[§] Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309, USA

^{||} IGPP, Scripps Institution of Oceanography, La Jolla, California 92093, USA

[¶] Department of Geological Sciences, University of Southern California, Los Angeles, California 90089, USA

Recent damaging earthquakes in California, including the 1971 San Fernando¹, 1983 Coalinga², 1987 Whittier Narrows³ and 1994 Northridge⁴ events, have drawn attention to thrust faults as both potentially hazardous seismic sources and as a mechanism for accommodating shortening in many regions of southern California. Consequently, many geological studies^{5,6} have concluded that thrust faults in Southern California pose the greatest seismic hazard, and also account for most of the estimated 5–7 mm yr⁻¹ of contraction across the greater Los Angeles metropolitan area^{7,8} indicated by Global Positioning System geodetic measurements⁹. Our study demonstrates, however, that less than 50% of the

geodetically observed contraction is accommodated on the principal thrust systems across the Los Angeles region. We integrate the most recent geological, geodetic and seismological data to assess the spatial distribution of strain across the Los Angeles metropolitan region. We then demonstrate that a significant component of seismic moment release and shortening in this region is accommodated by east–west crustal escape ‘extrusion’ along known strike-slip and oblique-slip faults.

The Los Angeles metropolitan region lies within a transitional zone where predominantly strike-slip rigid-block tectonics to the south gives way to east–west-trending folding and contractional faulting to the north (Fig. 1; ref. 10). The structural framework of this region is a product of polyphase deformation that includes Miocene extension and clockwise rotation, Pliocene contraction, and Plio-Quaternary transpression (oblique compression)^{11–13}. Many faults that are currently recognized as active show a complex history of movement; in some cases, varied senses of slip are evident¹⁴. Consequently, simple mechanical models of fault nucleation based on Coulomb fracture criteria do not readily explain the orientation of known faults with respect to presently observed maximum principal stresses (approximately north–south) defined by seismicity¹⁵ and borehole breakout studies¹⁶. Therefore, the activity and sense of slip on many active faults in the Los Angeles region are dictated by their pre-existing orientation relative to the current state of stress.

Recent studies show that contemporary crustal strain, regional north–south shortening and east–west extension in south California are being accommodated by rotations and by strike-slip, oblique-slip and conjugate faulting^{7,11–13,17,18}. Molnar¹⁹ compared the deformation within the western Transverse Ranges to extrusion tectonics in China; however, he focused on active, continuing rotations as a more accurate description of the deformation. Humphreys²⁰ proposed westward escape of the San Gabriel block, accommodated by left lateral faulting on its southern margin, as a mechanism to avoid convergence associated with the Big Bend of the San Andreas fault. However, recent kinematic and seismic hazard models for the Los Angeles metropolitan region have emphasized that most of the geodetically determined north–south shortening rates are accommodated on the principal thrust systems, such as the Sierra Madre, Elysian Park and Santa Monica faults, and have largely overlooked the role of strike-slip and oblique-slip faults as potentially hazardous seismic sources and as contractional accommodation structures. High rates of slip (3.8–5.5 mm yr⁻¹; refs 6, 21) have been inferred for the Sierra Madre–Cucamonga fault zone in part because of the substantial relief in its hanging wall and magnitude of deformation in Quaternary sediments on the foot wall²². However, geological observations west of the Cucamonga strand do not support high late Quaternary rates of slip. Our recent work shows that the rate of dip slip decreases from ~3–5 mm yr⁻¹ along the Cucamonga fault to ~1 mm yr⁻¹ for the central Sierra Madre fault zone and 1.5–3.0 mm yr⁻¹ for the San Fernando segment²³. The resulting shortening rate of ~1 mm yr⁻¹ along the central Sierra Madre fault zone is thus far less than is required to account for a large percentage of the 7–9 mm yr⁻¹ of north–south shortening across the Los Angeles region estimated from our analysis of the past three years of continuous Global Positioning System (GPS) measurements from the Southern California Integrated GPS Network (SCIGN)^{24,25}.

The San Jose, Raymond and other sinistral oblique-slip faults extend west-southwest away from the central section of the Sierra Madre fault zone where the slip rate is ~1 mm yr⁻¹. Their tectonic geomorphic expression²⁶ complements left lateral focal mechanisms and aftershock patterns^{27,28} from the 1988 and 1990 Upland earthquakes ($M_L = 4.6$ and $M_L = 5.2$) on the San Jose fault, and the 1988 Pasadena earthquake ($M_L = 4.9$) on the Raymond fault (M_L is Richter magnitude). The Cucamonga and San Fernando segments have significantly greater slip rates as they extend east and west of

[‡] Present address: Earth Consultants International, 2522 N. Santiago Blvd, Suite B, Orange, California 92667, USA.

the sinistral faults. This suggests that the conjugate strike-slip faults bounding the Pomona, Verdugo and Santa Monica blocks accommodate shortening across the Sierra Madre fault zone as the intervening crustal blocks rotate clockwise, and escape westward.

We assessed the strain distribution in the Los Angeles region using two independent methods; geology and geodesy (Figs 2 and 3,

Table 1). Analysis of site velocities estimated from the past three years of SCIGN continuous GPS data were used to calculate strain rates across structures in the Los Angeles region. Relative velocities were computed with respect to several different sites, and hence different perspectives, and synthesized to determine strain rates across known faults. In some cases the far-field strain across two or

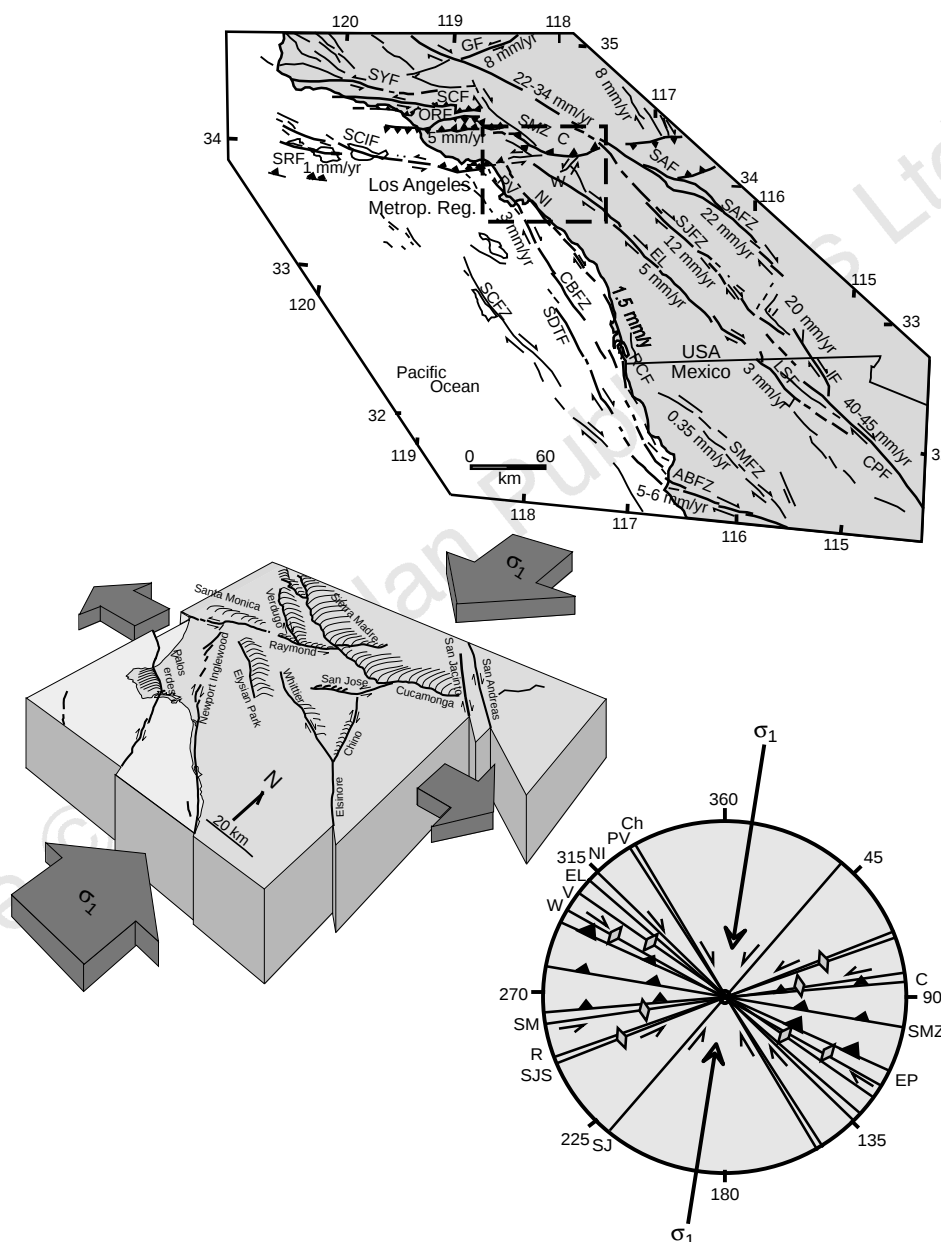


Figure 1 The main faults of southern California, northern Baja California, Mexico and the Los Angeles metropolitan region are shown on the regional map at the top of the figure. Arrows show relative lateral motion, barbs indicate thrust faulting and diamonds show oblique-slip faults. Published estimates of fault slip rates are shown in mm yr^{-1} . V, Verdugo; EP, Elysian Park; SJS, San Jose South; SJ, San Jose; Ch, Chino; SM, Santa Monica; R, Raymond; GF, Garlock fault; SYF, Santa Ynez fault; SCF, San Cayetano fault; ORF, Oak Ridge fault; SAF, San Andreas fault; SMZ, Sierra Madre; C, Cucamonga fault zone; SCIF, Santa Catalina Island fault, SRF Santa Rosa fault; PV, Palos Verdes fault; W, Whittier fault; NI, Newport Inglewood fault zone; CBFZ, Catalina Basin fault zone, SCFZ, San Clemente fault zone; SDTF, San Diego trough fault; SJFZ, San Jacinto fault zone; EL, Elsinor fault zone; RCFZ, Rose Canyon fault zone; IF, Imperial fault; LSF, Laguna Salada fault; SMFZ, San Miguel fault zone; ABFZ, Agua Blanca fault zone; CPF, Cerro Prieto fault. The schematic block (centre) and fault strike (bottom) diagrams show active faults in

the Los Angeles basin relative to the current tectonic stress regime. N-S compression forms strike and oblique-slip faults at moderate angles to the maximum principal stress and creates dominant compressional structures perpendicular to σ_1 (here σ_1 is the maximum principal stress direction). The spatial distribution of thrust, strike-slip and oblique-slip focal mechanisms also illustrates the structural complexity across the Los Angeles region and central Transverse Ranges. Microseismicity and mainshocks that define fault location, orientation and sense of slip agree with fault characteristics determined by palaeoseismic and geological studies. Fault plane solutions trending E-W indicate thrusting, whereas NW and NE solutions have predominantly strike-slip focal mechanisms²⁷⁻³¹. Left-lateral oblique slip faults such as the San Jose and Raymond form conjugate sets to associated deformation along the right-lateral oblique-slip Whittier and Verdugo fault zones.

Table 1 Measured and predicted slip rates

Fault	Ref.	Geological (mm yr ⁻¹)	Geodetic (mm yr ⁻¹)	Predicted (mm yr ⁻¹)
Blind thrusts and reverse				
West Sierra Madre	23	2.0 ± 1.0/-0.5	2.5 ± 1.7*	2.5 ± 0.5
Central Sierra Madre	23	1.0 ± 0.5	\$	
Cucamonga	23	3.0 ± 1.0	4.2 ± 1.5	
East Los Angeles	37	0.8 ± 0.2	\$	
Elysian Park	5	1.4 ± 0.2II	\$	
Northridge	38	1.6 ± 0.1/-0.2II		
Santa Monica/Hollywood	6, 39	1.0 ± 0.5/-0.6	3.5 ± 1.0	2.7 ± 1.0
Strike and oblique slip				
Verdugo		0.5†	2.1 ± 1.3	2.5 ± 1.0
Chino		0.5†	\$	1.5 ± 0.5
San Jose		0.5†	\$	1.5 ± 0.5
Clamshell-Sawpit		0.5†	\$	1.0 ± 0.5
Raymond		0.4†	\$	1.5 ± 0.5
Newport Inglewood	36	0.5 ± 0.5/-0.4		
Palos Verdes	33	2.8 ± 0.2/-0.2	2.9 ± 1.6‡	
Elsinore	35	4.9 ± 1.0/-0.5	4.9 ± 1.4	
San Gabriel	10	0.2 ± 0.5/-0.1		
Whittier	34	2.5 ± 0.5	\$	

Shown is a summary of slip rates determined by geological investigations^{5,6,33-39}, geodetically estimated cross fault strain rates, and our model predicted slip rates.

* Includes Northridge blind thrust and other structures in the San Fernando Valley.

† Published rates based on no data.

‡ Includes the Newport Inglewood fault zone.

\$ Estimated north-south contraction across structures in the northern Los Angeles region is $5.0 \pm 1.0 \text{ mm yr}^{-1}$ from our geodetic measurements. These include the Elysian Park system, and possibly unrecognized blind thrusts, and by associated structures of the central Sierra Madre, San Jose, Raymond, Whittier and Chino faults.

II Geological slip rate estimates for the Elysian Park and Northridge blind thrusts are based on two-dimensional balanced cross-section modelling that does not account for rotations or out of plane translation. However, without a dominant lateral component, the effects of out of plane motion on slip rate estimates are probably not substantial. Moreover, stratigraphic age errors that may result in an increase or decrease of two-dimensionally determined dip slip rates of 50% do not radically change our results.

more structures was used to reduce errors. Velocities of sites in the Los Angeles region relative to sites in the San Gabriel and Verdugo Mountains show 7–9 mm yr⁻¹ north-south contraction across the region, clockwise rotation, and east-west extrusion of blocks in the northern part of the study area (Fig. 2).

Comparison of the geological and geodetic strain distribution shows little deviation in north-south shortening across the western two-thirds of the study area for both geological and geodetic measurements (Fig. 3), demonstrating the conservation of north-south strain across multiple structures. Average north-south shortening using the geological data is $\sim 6.0 \text{ mm yr}^{-1}$ for the western two-thirds of the basin and increases to $\sim 8.0 \text{ mm yr}^{-1}$ where the Cucamonga fault intersects the central section of Sierra Madre fault zone. Similarly, the GPS results show $\sim 7.5 \text{ mm yr}^{-1}$ north-south contraction for the western two-thirds of the Los Angeles basin, which increases to 9.5 mm yr^{-1} near the Cucamonga fault. The increased strain budget associated with the Cucamonga fault can be explained by local strain partitioning from the San Andreas, San Jacinto and San Jose faults. Differential velocities ($4.0 \pm 1.3 \text{ mm yr}^{-1}$) between the San Gabriel block and the Perris block are accommodated across the Cucamonga fault zone as the restraining 'Big Bend' retards translation of the San Gabriel Mountains along the San Andreas fault. Coincident with the step in strain distribution between the Cucamonga fault and the slowly slipping central Sierra Madre fault, conjugate fault slip along the San Jose, Chino, Raymond and Verdugo faults effectively accommodate a component of north-south shortening by east-west block extrusion. The 1.5 mm yr^{-1} discrepancy between the total geodetic and geological rates probably represents the underestimation of

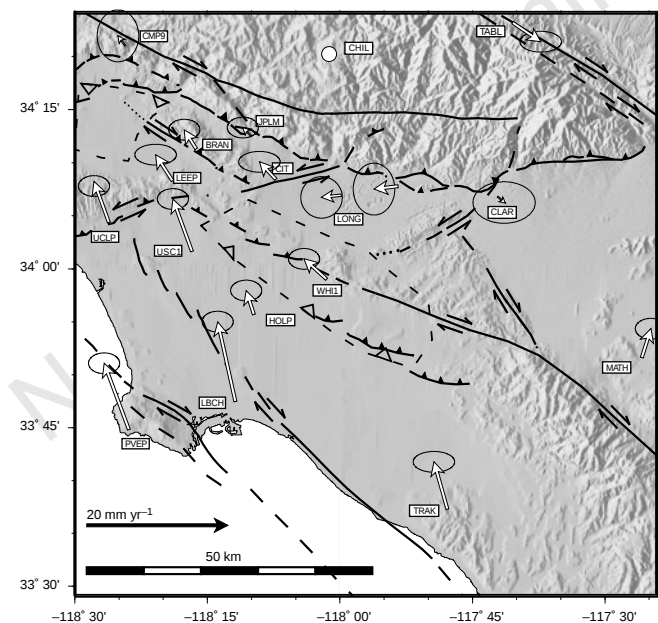
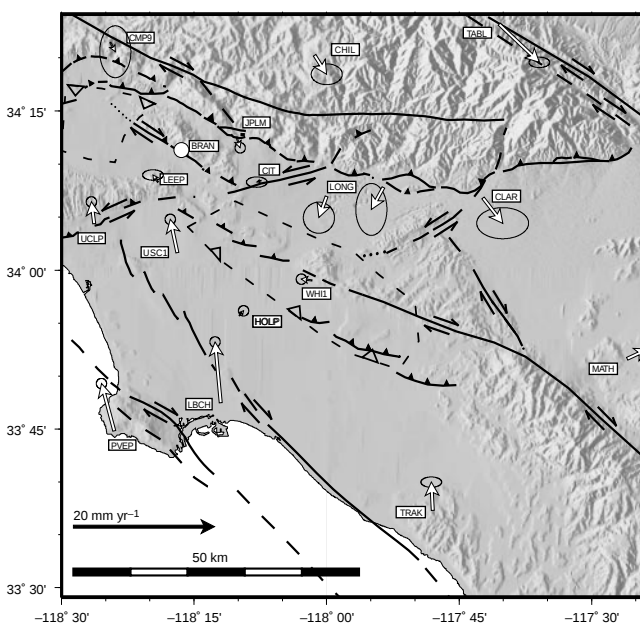


Figure 2 Geodetic velocity field relative to sites CHIL and BRAN in the San Gabriel and Verdugo mountains. Error ellipses show 95% confidence assuming a fractal white noise model for the daily position estimates (the ellipses are scaled from the white noise estimate by a factor of three to account for the coloured noise³²). From the perspective of CHIL (left panel), we see (1) westward escape of the Pomona block which is bound on the north by Sierra Madre fault zone, and left-lateral oblique-slip San Jose and Raymond faults; note there is little contraction across the slowly slipping central Sierra Madre faults zone; (2) 4–7 mm yr⁻¹ NNE motion of MATH on the Perris block, most of which is accommodated on the Cucamonga fault zone; (3) E-W oblique extension between LEEP on the Santa Monica block and MATH suggesting E-W extrusion across the northern Los Angeles Metropolitan region; (4) clockwise rotation of vectors from east to west across the region; (5) 10 mm yr⁻¹ of NNW dextral shear of PVEP and 6–7 mm yr⁻¹ of contrac-



tion of PVEP with most of the contraction accommodated between USC1 and JPLM, suggesting that significant shortening is accommodated across the Raymond-Santa Monica/Verdugo conjugate faults and associated structures of the Elysian Park blind thrust; (6) northwest motion of BRAN and CIT on the Verdugo block with little convergence across the slowly slipping central Sierra Madre fault zone. Geodetic velocities with respect to BRAN (right panel) in the Verdugo Mountains show (1) 3–5 mm yr⁻¹ eastward motion of MATH on the Perris block suggesting east-west extrusion across multiple conjugate fault systems in the northern Los Angeles region; (2) 4–6 mm yr⁻¹ E-W extension between LEEP and MATH; (3) clockwise rotation of vectors from north to south across the region; (4) 1–3 mm yr⁻¹ of NW motion of LEEP, suggesting dextral shear across the Verdugo fault and westward escape of the Santa Monica block.

geological slip rates for faults in the study area, or may also suggest the presence of unrecognized blind thrusts. Conversely, overlapping errors in slip and strain rates could resolve this discrepancy.

Comparison of the east–west distribution of strain is not as easily interpreted as the results of north–south shortening. Geologically determined rates of east–west motion sum to $\sim 5.5 \text{ mm yr}^{-1}$ for the southern Los Angeles region, and decrease abruptly to as low as 1 mm yr^{-1} farther north near the Sierra Madre–Cucamonga fault zone. These results are inconsistent with rates derived from geodetic studies which are $\sim 6 \text{ mm yr}^{-1}$ across the entire area of study. Assuming that the geodetic measurements are accurate, the discrepancy in east–west motion between the geodetic and geological results indicates that slip rates for the Verdugo, Raymond, San Jose,

and Chino faults may be higher than previously estimated. Increasing the sum east–west geological rates into closer agreement with the geodetic rates yields slip rate estimates as follows; $1.0\text{--}3.0 \text{ mm yr}^{-1}$ of right lateral slip and $0.5\text{--}1.5 \text{ mm yr}^{-1}$ of shortening for the Verdugo fault; $1.0\text{--}2.0 \text{ mm yr}^{-1}$ of left lateral slip along the Raymond fault; $1.0\text{--}2.0 \text{ mm yr}^{-1}$ of left-oblique slip the San Jose fault; and $1.0\text{--}2.0 \text{ mm yr}^{-1}$ right lateral slip for the Chino faults (Table 1). These revised slip rate estimates bring the north–south geodetic and geological components into very close agreement, and also result in a change in the overall perceived seismic hazard.

In a synoptic study of the hazard posed by faults within the greater Los Angeles metropolitan area, Dolan *et al.*⁶ used geological estimates of slip rates to infer average recurrence intervals for

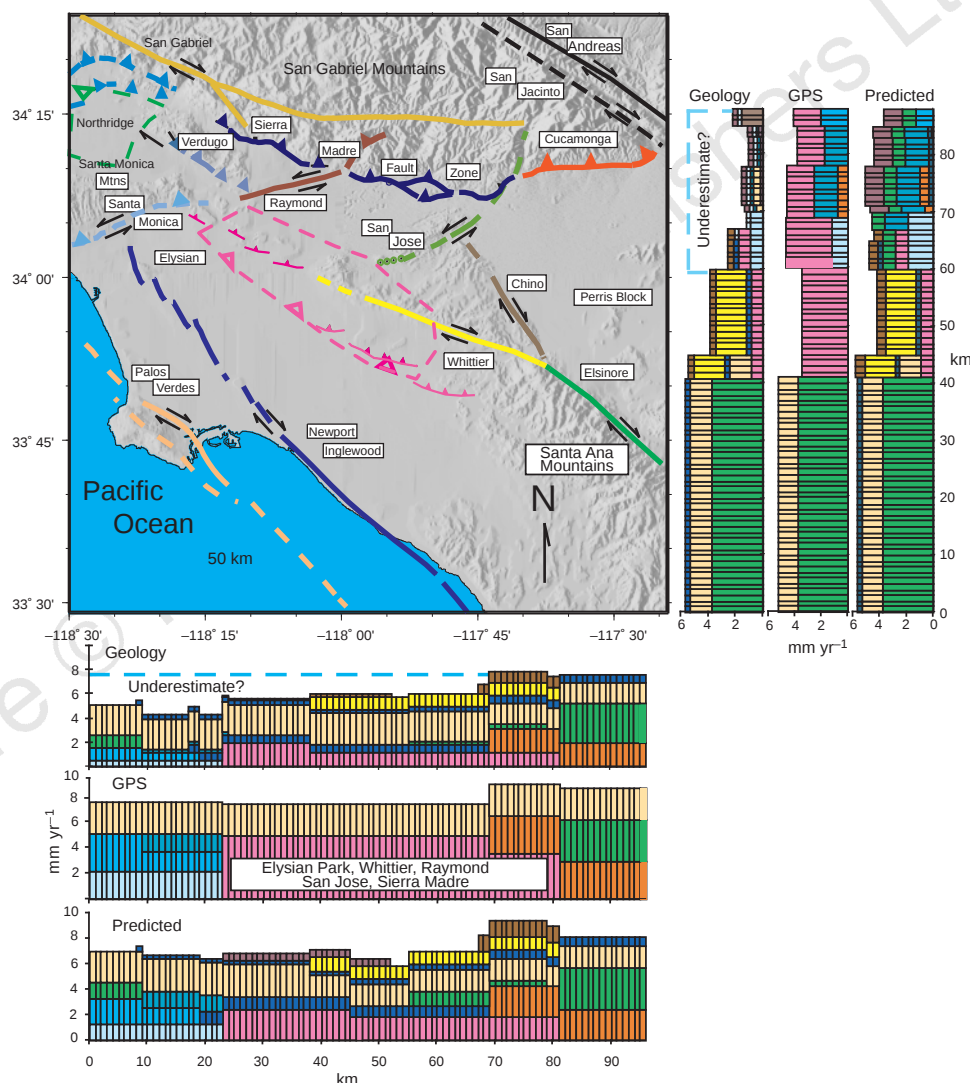


Figure 3 North–south and east–west cumulative distribution of strain across the Los Angeles metropolitan region based on geological data, geodetic data and our predicted slip rates (Table 1). Geological data were compiled from previous studies, data in the Southern California Earthquake Center (SCEC) Phase 3 report¹⁰ and our recent work along the Sierra Madre–Cucamonga fault zone. Geodetic cross fault strain rates were estimated by fault parallel relative site velocities, baseline extension and baseline contraction between sites. Fault geometry (dip, strike, rake), sense of slip, and slip rate were resolved into horizontal N–S and E–W components along the length of each fault (that is, Elysian fault (green) slips at 5.0 mm yr^{-1} on a vertical fault that strikes 310° : N–S motion = $(5.0 \text{ mm yr}^{-1}) \times [\cos(360-310)] = 3.2 \text{ mm yr}^{-1}$ for the eastern 15 km of the study area (the Whittier and Chino faults accommodate the Elysian component farther to the west); E–W motion = $(5.0 \text{ mm yr}^{-1}) \times [\sin(360-310)] =$

3.8 mm yr^{-1} for the southern 40 km of the study area). Fault characteristic variations for individual segments were accounted for (that is, strike change producing increase/decrease in contractional component). The region was then divided by a N–S and E–W grid at 1 km increments. The location and magnitude of strain for each km^2 block were then projected onto N–S and E–W cross-sections to illustrate spatial patterns of crustal strain. The lower graphs represent cumulative N–S shortening from east to west across the region shown. The graphs on the right represent cumulative E–W motion from north to south. Colour of blocks correspond to the fault of the same colour. Total geodetic sums of N–S contraction and E–W extrusion across the Elysian Park, Whittier, Raymond, San Jose and Sierra Madre faults are shown as pink blocks and were calculated from baseline extension and contraction between multiple sites.

different earthquakes occurring on faults throughout the region. The faults we discuss here make up only a subset of the faults studied by Dolan *et al.*, who examined a larger region. Nevertheless, a comparison between the two estimates is a useful measure of how our slip rates revise the estimate of the seismic hazard posed by these urban faults. If we compare just the area of overlap between the two studies, our slip rates yield a total moment accrual rate that is only 65% to 85% of the rate inferred by Dolan *et al.*, closer to that observed historically. Our slower central Sierra Madre fault zone slip rate is only partially balanced by the addition of the Chino fault, higher slip rates for the Raymond and San Jose faults, and the relatively high slip rate for the Verdugo fault. If we consider the entire area described by Dolan *et al.*, however, the net effect of our proposed slip rates on the seismic hazard indicates a decrease in moment accrual rate of only ~2% to 7%.

We have combined geodetic and geological data in an attempt to improve seismic risk assessment for the greater Los Angeles metropolitan region. We interpret the results of our analysis to reflect the important role of conjugate strike-slip faults in this region which apparently accommodate over 50% of the geodetically observed north-south shortening. This implies that calculations of moment release may have been previously overestimated for reverse and blind thrust sources, which presented a greater perceived risk of future damaging earthquakes. Thus, our observations and model suggest that the Verdugo, Raymond, San Jose and Chino faults may play a larger role in moment release in Los Angeles region, in spite of their relatively short lengths and lack of large historical earthquakes. This implies that the distribution of seismic moment release is likely to occur as smaller-magnitude events (that is, on faults with smaller surface areas) punctuated by less-frequent, larger-magnitude earthquakes. □

Received 19 May 1997; accepted 6 April 1998.

- Oakeshott, G. (ed.) *San Fernando, California Earthquake of 9 February 1971* (Bull. 196, California Division of Mines and Geology, Sacramento, CA, 1975).
- Namson, J. & Davis, T. Seismically active fold and thrust belt in the San Joaquin Valley, central California. *Geol. Soc. Am. Bull.* **100**, 257–273 (1988).
- Hauksson, E. & Jones, L. The Whittier Narrows earthquake sequence in Los Angeles, California: seismological and tectonic analysis. *J. Geophys. Res.* **94**, 9569–9590 (1989).
- Hauksson, E., Jones, L. & Hutton, K. The 1994 Northridge earthquake sequence in California. Seismological and tectonic aspects. *J. Geophys. Res.* **100**, 12335–12355 (1995).
- Shaw, J. H. & Suppe, J. Earthquake hazards of active blind thrust faults under the central Los Angeles basin, California. *J. Geophys. Res.* **101**, 8623–8642 (1996).
- Dolan, J. F. *et al.* Prospects for larger or more frequent earthquakes in the Los Angeles metropolitan region. *Science* **267**, 199–205 (1995).
- Feigl, K. L. *et al.* Space geodetic measurement of crustal deformation in central and southern California, 1984–1992. *J. Geophys. Res.* **98**, 21677–21712 (1993).
- Shen, Z.-K. *et al.* Crustal deformation across and beyond the Los Angeles basin from geodetic measurements. *J. Geophys. Res.* **101**, 27957–27980 (1996).
- Kleusberg, A. & Teunissen, J. G. *GPS for Geodesy* (Lecture Notes in Earth Sciences, Springer, 1996).
- Phase III Report (Southern California Earthquake Center, Univ. Southern California, Los Angeles, 1996).
- Luyendyk, B. A model for Neogene crustal rotations, transtension, and transpression in southern California. *Geol. Soc. Am. Bull.* **103**, 1528–1536 (1991).
- Wright, T. L. in *Active Margin Basins* (ed. Biddle, K. T.) 35–134 (Mem. 52, Am. Assoc. Petroleum Geologists, Tulsa, OK, 1991).
- Dickinson, W. Kinematics of transrotational tectonism in the California Transverse Ranges and its contribution to cumulative slip along the San Andreas transform fault system. *Geol. Soc. Am. Spec. Pap.* **305**, 1–50 (1996).
- Tsutsumi, H. & Yeats, R. Late Cenozoic structural geology of the 1971 San Fernando and 1994 Northridge earthquake zones in the northern San Fernando Valley, California. *Eos* **76**, F361 (1995).
- Hauksson, E. Earthquakes, faulting, and stress in the Los Angeles Basin. *J. Geophys. Res.* **95**, 15365–15394 (1990).
- Mount, S. & Suppe, J. Present-day stress orientations adjacent to active strike-slip faults, California and Sumatra. *J. Geophys. Res.* **97**, 11995–12103 (1992).
- Nicholson, C., Sorlien, C. C., Atwater, T., Crowell, J. C. & Luyendyk, B. P. Microplate capture, rotation of the western Transverse Ranges, and initiation of the San Andreas transform as a low-angle fault system. *Geology, Geol. Soc. Am.* **22**, 491–495 (1994).
- Donnellan, A., Hager, B. H., King, R. W. & Herring, T. A. Geodetic measurement of deformation in the Ventura basin region, southern California. *J. Geophys. Res.* **98**, 21727–21739 (1993).
- Molnar, P. *Final report to the Southern California Earthquake Center* (Univ. Southern California, Los Angeles, 1992).
- Humphreys, E. Tectonics of the greater Los Angeles region: Implications for the lower crust and upper mantle. *Eos* **76**, F347 (1995).
- Morton, M. & Matti, J. in *Recent Reverse Faulting in the Transverse Ranges, California* (eds Morton, D. M. & Yerkes, R. E.) 179–203 (US Geol. Surv. Prof. Pap. 1339, Washington DC, 1987).
- Ehlig, P. in *San Fernando, California Earthquake of 9 February 1971* (ed. Oakeshott, G. B.) 7–18 (Bull. 196, California Division of Mines and Geology, Sacramento, CA, 1975).
- Walls, C., Rockwell, T. K., Pfanner, J., Bornyaz, M. & Lindvall, S. Uplift gradient along the Sierra Madre-Cucamonga fault zone, Los Angeles, California. *Geol. Soc. Am. Abstr. Cordilleran Sect. Mtg* **29**, A-72 (1997).

- Bock, Y. Southern California Permanent GPS Geodetic Array. Continuous measurements of crustal deformation between the 1992 Landers and 1994 Northridge earthquakes. *J. Geophys. Res.* **102**, 18013–18033 (1997).
- Bock, Y. & Williams, S. Integrated satellite interferometry in southern California. *Eos* **78**, 293299–293300 (1997).
- Crook, R. J., Allen, C. R., Kamb, G., Payne, C. M. & Proctor, R. J. in *Recent Reverse Faulting in the Transverse Ranges, California* (eds Morton, D. M. & Yerkes, R. E.) 7–26 (US Geol. Surv. Prof. Pap. 1339, Washington DC, 1987).
- Hauksson, E. & Jones, L. The 1988 and 1990 Upland earthquakes. Left-lateral faulting adjacent to the central Transverse Ranges. *J. Geophys. Res.* **96**, 8143–8165 (1991).
- Jones, L. M., Sieh, K., Hauksson, E. & Hutton, L. K. The 3 December 1988 Pasadena, California earthquake. Evidence for strike-slip motion on the Raymond fault. *Bull. Seismol. Soc. Am.* **80**, 474–482 (1990).
- Penchman, J. in *Recent Reverse Faulting in the Transverse Ranges, California* (eds Morton, D. M. & Yerkes, R. E.) 97–112 (US Geol. Surv. Prof. Pap. 1339, Washington DC, 1987).
- Press, F. & Allen, C. Patterns of seismic release in the southern California region. *J. Geophys. Res.* **100**, 6421–6430 (1995).
- Hutton, L. K., Jones, L. M. & Given, D. D. in *Neotectonics of North America* (eds Slemmons, D. B., Engdahl, E. R., Zoback, M. D. & Blackwell, D. D.) 133–152 (Geol. Soc. Am., Boulder, CO, 1991).
- Zhang, J. *et al.* Southern California Permanent GPS Geodetic Array. Error analysis of daily position estimates and site velocities. *J. Geophys. Res.* **102**, 18035–18055 (1997).
- McNeilan, T. W., Rockwell, T. K. & Resnick, G. S. Style and rate of Holocene slip, Palos Verdes fault, southern California. *J. Geophys. Res.* **101**, 8317–8334 (1996).
- Rockwell, T., Gath, E. M. & Gonzalez, T. Sense and rate of slip on the Whittier fault zone, eastern Los Angeles Basin, CA. *Assoc. Eng. Geol. Abstr.* **35**, 679 (1992).
- Bergmann, M. & Rockwell, T. Preliminary assessment of the late Holocene slip rate for the Wildomar fault in Murrieta, California. *Assoc. Eng. Geol. Abstr. Programs* **36**, 43 (1993).
- Grant, L. B., Waggoner, J. T., Rockwell, T. & von Stein, C. Paleoseismicity of the north branch of the Newport-Inglewood fault zone in Huntington Beach, California, from cone penetrometer test data. *Bull. Seismol. Soc. Am.* **87**, 277–293 (1997).
- Oskin, M. & Sieh, K. The Elysian Park anticlinorium: Surficial evidence of an active blind reverse fault beneath downtown Los Angeles. *Geol. Soc. Am. Abstr. Cordilleran Sect. Mtg* **30**, A-57 (1998).
- Yeats, R. S. & Huftile, G. J. The Oak Ridge fault system and the 1994 Northridge earthquake. *Nature* **373**, 418–420 (1995).
- Dolan, J. F., Sieh, K., Rockwell, T. K., Gupitill, P. & Miller, G. Active tectonics, paleoseismology, and seismic hazards of the Hollywood fault, northern Los Angeles basin, California. *Geol. Soc. Am. Bull.* **109**, 1595–1616 (1997).

Acknowledgements. We thank J. Shaw, C. Nicholson and R. Yeats for criticisms and suggestions, and the Southern California Integrated GPS Network and International GPS Service for continuous GPS data. This work was supported by the Southern California Earthquake Center.

Correspondence and requests for materials should be addressed to C. Walls (e-mail: cwalls@geology.sdsu.edu).

Evidence from the asymmetry of fast-spreading ridges that the axial topographic high is due to extensional stresses

Michael A. Eberle* & Donald W. Forsyth

Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

Along fast-spreading mid-ocean ridges such as the East Pacific Rise, there is an axial topographic high, 5–20 km wide, which stands 200–400 m above the background slope caused by thermally induced seafloor subsidence. There are also smaller topographic lows flanking the axial high along most of the East Pacific Rise from 20° S to 15° N. The existence of these lows is predicted by models of the origin of the axial high. One model postulates that the axial high is created by buoyant uplift from a narrow zone of concentrated partial melt extending tens of kilometres down into the mantle^{1–4}. Another model requires no such buoyant zone, suggesting instead that the axial high is generated by dynamic, extensional stresses in the lithosphere and shallow asthenosphere⁵. Here we show that the observed asymmetry of the flanking lows can be used to distinguish between these two proposed mechanisms. Although either model can be adapted to match the asymmetry on individual profiles, the along-axis variation in degree of symmetry favours the model of dynamic, extensional stresses for the origin of the axial high and its flanking lows.

* Present address: Laboratoire de Géophysique et Planétologie, Faculté des Sciences, 2 rue de la Houssinière, 44072 Nantes, France.

Creating the axial high by buoyant uplift from a narrow, low-density conduit containing partial melt extending tens of kilometres vertically down into the mantle¹⁻⁴ (Fig. 1a) requires upwelling beneath the ridge to be highly focused by dynamic feedback mechanisms. A narrow, buoyant channel is predicted by viscous-flow models that incorporate melt retention buoyancy, temperature-dependent viscosity and reduced viscosity where melt is present⁶⁻⁹. The extensional-stress model (Fig. 1b) requires that the horizontal normal stresses within the crust deepen near the ridge, creating moments that deflect the axial lithosphere upwards⁵. This

model has no *a priori* requirements of either focused or passive upwelling, or a narrow or broad melt zone, but does require that the effective, time-averaged stresses in the near-axis, uppermost crustal lithosphere are low compared to the underlying viscous crust, presumably by release of extensional stresses by dyking. Previously published versions of these models¹⁻⁵ were symmetric about the axis and could match only the symmetric component of axial topography. Near-axis topography along one of the best-surveyed parts of the East Pacific Rise (EPR), however, is clearly asymmetric, with a well developed, flanking low only on the west side (Fig. 2a).

The existence of asymmetric flanking lows clearly requires some asymmetry in crustal or mantle properties. In region A in Fig. 3 (the location of the averaged profiles shown in Fig. 2), there are abundant indications that such variations are present, including asymmetries in subsidence rates, absolute plate motion rates, spreading or accretion rates, seismic anisotropy, seamount abundance, and mantle shear wave velocity. We tested which properties might control the asymmetry of the axial topography by perturbing the symmetric models. Introducing asymmetric accretion rates or asymmetric lithospheric thickening had little effect in either the buoyant or extensional stress models. Asymmetry of the axial magma chamber in the extensional stress models did not produce enough asymmetry in the predicted topography to match the observations, nor did asymmetry in viscosity in the buoyant model.

In the buoyant model, we discovered that tilting the conduit produces an asymmetric low. A conduit tilted $\sim 8^\circ$ towards the east (Fig. 1a) can match the observed asymmetric topography along the 15° – 19° S section of the EPR (dashed line in Fig. 2a). Asymmetry could also be produced by broadening the conduit on one side, but

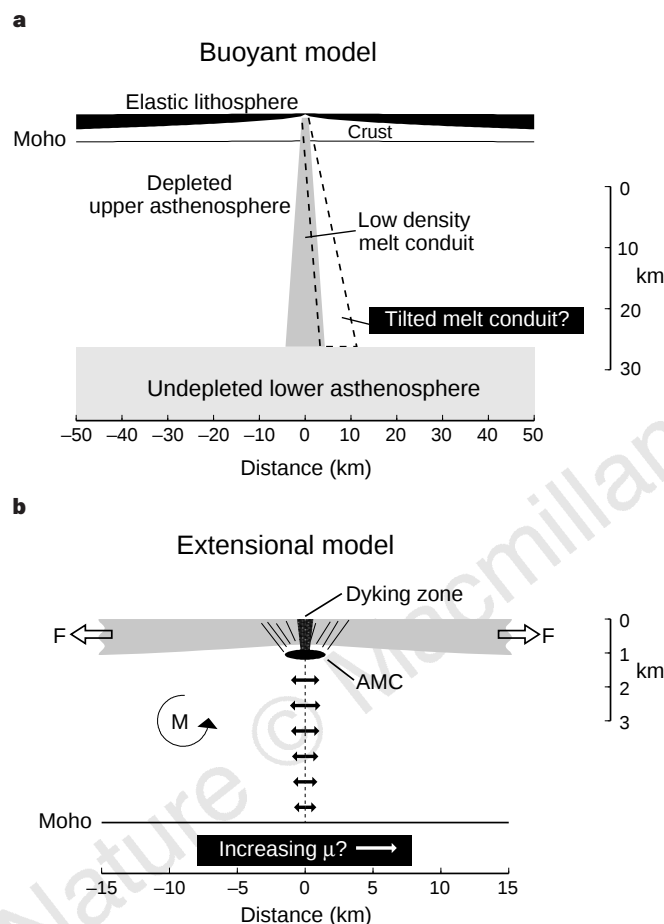


Figure 1 Two mechanisms that can produce an asymmetric axial high and flanking lows at the fast-spreading East Pacific Rise. **a**, Buoyancy forces in a zone of partial melt extending tens of km into the mantle lift up the overlying seafloor. When the viscosity in the melt conduit is much less than in the surrounding asthenosphere⁴, the upwards flow and buoyancy forces are focused into a narrow channel, producing a narrow topographic high. Flexural response of the plate to these uplifting forces creates flanking lows. If the base of the conduit is offset ~ 7 km to the east (dashed line), this model produces the asymmetric topography shown in Fig. 2a (short dashed line). **b**, Dynamic moments caused by extensional stresses may deflect the axial lithosphere upwards, creating the narrow axial high. At fast-spreading ridges, the lithosphere is thin at the axis and thickens slowly with distance away from the axis. A steady-state, axial magma chamber (AMC) may provide a nearly stress-free boundary and stresses in the lithosphere, near the axis, may be relieved by injecting dykes into the plate. In this scenario⁵, extensional stresses will be supported mostly within the lithosphere except at the axis, where they will be supported in the crust and mantle below the AMC. Near the ridge, the depth of the average extensional forces, $\bar{F}$, will decrease away from the axis, creating internal moments, M , within the crust. These moments are balanced by the topographic load of the axial high. When the viscosity in the crust and mantle decreases by about a factor of 3 from east to west (left) over a distance of 100 km, this model can produce asymmetric topography similar to that shown in Fig. 2a.

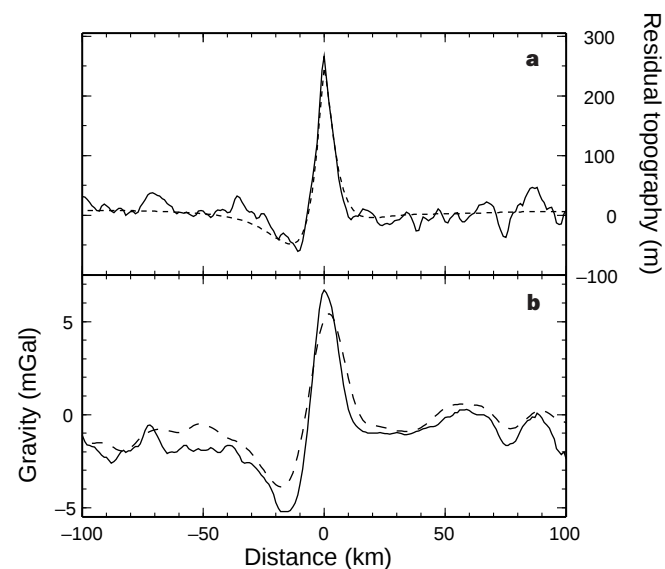


Figure 2 Characteristic profiles of residual topography and gravity of the EPR between 15° and 19° S. Across-axis profiles of the residual bathymetry and gravity are stacked and averaged to produce the characteristic profiles. **a**, Residual topography is calculated by subtracting from the observed topography, the predicted topography of a passive-flow, thermal model that matches the subsidence more than 8 km away from the axis to the east and west. Characteristic residual bathymetry (solid line) shows an axial high ~ 270 m shallower than the surrounding residual topography, with a width of ~ 15 – 20 km and a flanking low to the west of ~ 50 – 70 m. Dashed line shows the residual topography predicted by the buoyant model when the conduit tilts about 8° to the east. The extensional model can produce an equally good fit when there is a factor of 3 decrease in viscosity of the crust and mantle layers from east to west. **b**, Characteristic profiles of the shipboard (solid line) and satellite (dashed line) free-air gravity show a high and low corresponding to the bathymetric high and low.

it is difficult to match the amplitude of the flanking low on the west side unless the western edge of the conduit also slants to the east. Tilting of the conduit could be caused by migration of the ridge axis to the west relative to the underlying mantle. At this latitude on the EPR or Nazca/Pacific plate boundary, the ridge is migrating to the west at about 35 km Myr^{-1} in the hotspot coordinate frame, because the Pacific plate is moving twice as fast to the west as the Nazca plate is moving to the east¹⁰. Ridge migration is further enhanced by more rapid spreading or accretion of new seafloor on the Nazca plate than the Pacific plate¹¹. Viewed in a coordinate system fixed on the EPR, the deeper mantle is moving to the east, tending to displace upwelling particles or partial melt zones to the east of the spreading centre.

In the extensional-stress model, the axial high and flanking low are created by stresses within the lithosphere caused by separation of the lithospheric plates and flow in the underlying viscous asthenosphere (Fig. 1b). Our numerical experiments show that an effective way to match the asymmetric low and its accompanying gravity anomaly in this model is to introduce a decrease in viscosity in the crustal and mantle asthenosphere from east to west (Fig. 4). Although the specific model is non-unique, we find an excellent fit to the data, with viscosity in the crust and mantle decreasing by a factor of 3 from east to west over a distance of 100 km centred on the ridge axis. At temperatures where the crust and mantle deform viscously, a lateral viscosity gradient of this magnitude could be caused by average horizontal temperature gradients of approximately $0.3^\circ\text{C km}^{-1}$. This increase in temperature to the west is consistent with the observation that the western flank of the southern EPR subsides more slowly than the eastern flank^{4,12}. A

thermal gradient in the mantle of $\sim 0.1^\circ\text{C km}^{-1}$ can explain the long-lived asymmetric subsidence in this region¹², but a larger gradient may better fit the asymmetric subsidence within $\sim 150 \text{ km}$ of the ridge⁴. Also, the phase velocities of seismic Rayleigh waves are slower to the west, indicating that shear velocities are lower and temperatures are probably higher in the upper 100 km of the mantle beneath the Pacific plate near the EPR¹³. Other asymmetries could be responsible for viscosity variations of this magnitude, such as changes in the amount of melt or volatiles present, but the slower subsidence to the west implies that there are thermal gradients in the mantle.

Our calculations suggest that we should be able to distinguish between the buoyant model and the extensional-stress model by comparing variations in the degree and sense of asymmetry to the rate of migration of the ridge (tilting of the melt conduit) and to the asymmetry in subsidence (temperature and viscosity gradient). Because topographic coverage is incomplete and the flanking lows are more easily detected in the free-air anomaly than in the residual topography, we employ free-air anomaly maps derived from satellite altimetry¹⁴, with nearly uniform coverage to assess the degree of asymmetry of the flanking low as a function of latitude along the EPR. The same features are visible in the satellite data as in the shipboard-derived free-air anomalies (Fig. 2b). Although short-wavelength components of the satellite-gravity fields may be underestimated and there may be ringing effects from the filtering of the fields, the asymmetry of the lows, the amplitudes relative to the gravity signature of the axial high, and the agreement with shipboard data, preclude the possibility that the flanking lows visible in the satellite gravity are simply an artefact.

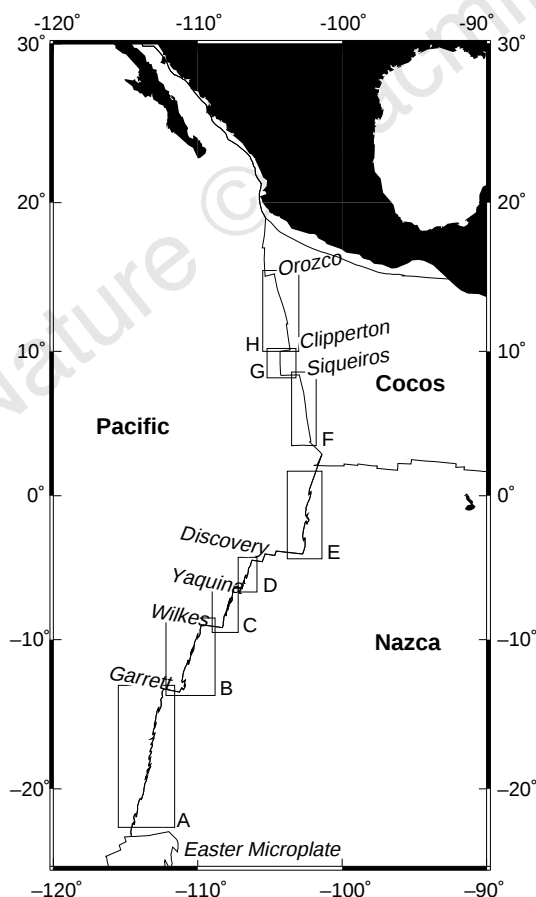


Figure 3 The East Pacific Rise from the Easter Microplate to North America. Labelled boxes show the eight regions examined in this study. Major transform offsets between the regions are labelled in italics. Figure 5c shows the ridge migration rate within each region in the fixed, hotspot reference frame.

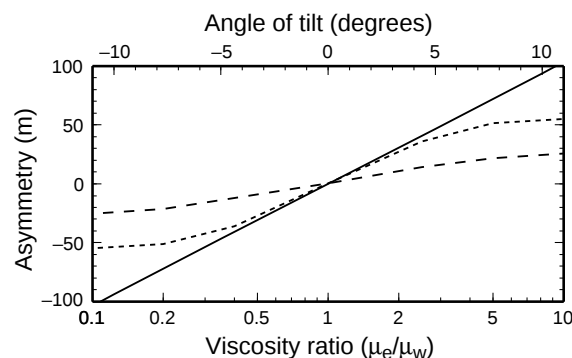


Figure 4 Degree of topographic asymmetry as a function of asymmetry in the models. Topographic asymmetry is the largest absolute difference between the predicted topography at equal distances from the axis on the two sides of the ridge. Positive values indicate that the flanking low is more pronounced (deeper) to the west. Solid line gives the asymmetry for varying angles of conduit tilt in the buoyant model. Positive tilt angles indicate that the base of the conduit is offset to the east. When the conduit is tilted, more of the buoyancy forces are transmitted vertically to one side of the ridge, which tends to eliminate the flexural depression that flanks the axial uplift on that side of the ridge. The best-fitting, buoyant model has an asymmetry of 62 m (short dashed line in Fig. 2a). Dashed lines give the asymmetry when there are viscosity gradients in the crust (long-dashed line) or mantle (short-dashed line) in the extensional model. Viscosity ratio describes the total change in viscosity from west to east when a linear gradient is introduced extending from 50 km west of the axis to 50 km east. When there is asymmetry in viscosity, there is more effective coupling of the plate to the asthenosphere on the high viscosity side. This coupling transfers some of the stress to greater depth, reducing the near-axis moment on that side of the ridge, and drags some asthenospheric material from the low-viscosity region to the high-viscosity region, reducing the pressure at the base of the plate on the low-viscosity side. Although different combinations of viscosity gradients in crust and mantle are possible, an excellent fit to the topography is obtained when the total change is a factor of 3 in both the crust and mantle.

To measure the asymmetry, we calculate the area of the flanking lows in the average free-air gravity profile (Fig. 5a) for each of eight regions along the EPR (boxes A–H in Fig. 3). The greatest asymmetry is in the 13°–23° S section (Fig. 5b). The degree of asymmetry decreases to the north, but remains significant up to at least the Yaquina and perhaps to the Discovery fracture zone at ~4° S. To the north of this point, there are variations in amplitude of the flanking lows, but no asymmetry is found.

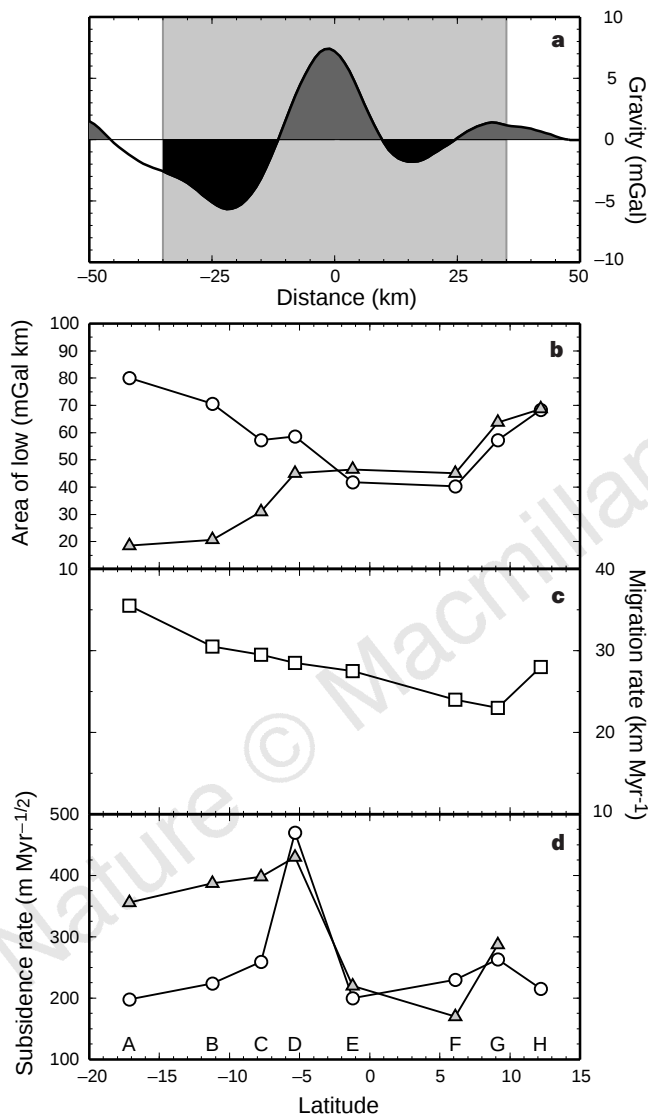


Figure 5 Along-axis variations in properties of the East Pacific Rise. **a**, A characteristic free-air gravity profile is calculated for each region in Fig. 3 by stacking values from a 2.5° by 2.5° (lat. by long.) grid of the satellite gravity¹⁴ and removing the mean. The integrated areas below the mean on both sides of the axis (dark-shaded regions) out to 35 km provide a parameterized description of the flanking gravity lows. Given the small size of the anomalies and the amplitude of the noise, we find that the integrated anomaly is a more stable estimate than its maximum value. **b**, Area of the flanking gravity low to the east (triangles) and west (circles) of the ridge for each region shown in Fig. 3. **c**, Motion of the ridge towards the west relative to hotspots within each region. Migration of the ridge with respect to the mantle may cause the top of a melt conduit to be offset from the base, creating asymmetric topography in the buoyant model. **d**, Subsidence rates to the east (triangles) and west (circles) of the EPR within each region. Asymmetric subsidence may be indicative of thermal gradients in the crust and mantle. These gradients will produce asymmetric topography in the extensional model.

Because the northern part of the EPR is also migrating rapidly in the hotspot coordinate frame (Fig. 5c), it is unlikely that the change from asymmetric to symmetric flanking topography (Fig. 5b) is caused by a corresponding change from a tilted to a vertical conduit. In contrast, there is an excellent correlation between asymmetry in subsidence of the seafloor and asymmetry in the flanking low (Fig. 5b, d). South of the Yaquina Fracture Zone, the Pacific plate subsides much more slowly than the Nazca plate and the flanking low is largely confined to the west side of the ridge. Further north, although subsidence is also slow in some sections, it is nearly symmetric and the flanking low is equally well developed on both sides of the ridge. Thus, where seafloor subsidence suggests an across-axis gradient in mantle temperature and viscosity, there is a systematic asymmetry in amplitude of the flanking low. Because there is agreement between the asymmetry in subsidence rate and the flanking low along the EPR and no agreement between the migration rate of the ridge and the degree of asymmetry of the flanking low (Fig. 5b–d), we favour the extensional-stress model over the buoyant model for the origin of the axial high.

It has been proposed that the axial topographic high is compensated by buoyancy forces extending tens of kilometres into the mantle^{1–4} (Fig. 1a). For this model, upwelling must be highly focused and approximately two-dimensional beneath fast-spreading ridges to produce the narrow melt conduit and linear axial high observed along most of the EPR. The extensional model favoured in our study does not have any *a priori* requirements on the asthenospheric flow or geometry of the melt zone, but does require that the extensional stresses deepen towards the ridge. At slow spreading rates, the cooling plates thicken more rapidly with distance from the axis, and the characteristic median valley may form because this thickening causes the extensional stresses to deepen away from the ridge¹⁵. Thus, as the extensional-stress model may be able to explain the along-axis variations of the axial topography observed along the EPR, whereas the buoyant model cannot, the axial topography (that is, axial high versus median valley) may be indicative of the rheology of the shallow crust and mantle, rather than of the deep flow field and geometry of the zone of partial melt in the sub-axial asthenosphere. □

Received 6 November 1997; accepted 29 April 1998.

- Wilson, D. S. Focused upwelling beneath mid-ocean ridges: Evidence from seamount formation and isostatic compensation of topography. *Earth Planet. Sci. Lett.* **113**, 41–55 (1992).
- Wang, X. & Cochran, J. R. Gravity anomalies, isostasy, and mantle flow at the East Pacific Rise crest. *J. Geophys. Res.* **98**, 19505–19531 (1993).
- Magde, L. S. *et al.* Crustal and upper mantle contribution to the axial gravity anomaly at the southern East Pacific Rise. *J. Geophys. Res.* **100**, 3747–3766 (1995).
- Eberle, M. A., Forsyth, D. W. & Parmentier, E. M. Constraints on a buoyant model for the formation of the axial topographic high on the East Pacific Rise. *J. Geophys. Res.* **103**, 12291–12307 (1998).
- Eberle, M. A. & Forsyth, D. W. An alternative, dynamic model of the axial topographic high at fast-spreading ridges. *J. Geophys. Res.* **103**, 12309–12320 (1998).
- Buck, W. R. & Su, W. Focused mantle upwelling below mid-ocean ridges due to feedback between viscosity and melting. *Geophys. Res. Lett.* **16**, 641–644 (1989).
- Scott, D. R. & Stevenson, D. J. A self-consistent model of melting, magma migration and buoyancy-driven circulation beneath mid-ocean ridges. *J. Geophys. Res.* **94**, 2973–2988 (1989).
- Sotin, C. & Parmentier, E. M. Dynamical consequences of compositional and thermal density stratification beneath spreading centers. *Geophys. Res. Lett.* **16**, 835–838 (1989).
- Su, W. & Buck, W. R. Buoyancy effects on mantle flow under mid-ocean ridges. *J. Geophys. Res.* **98**, 12191–12205 (1993).
- Gripp, A. E. & Gordon, R. G. Current plate velocities relative to the hotspots incorporating the NUVEL-1 global plate motion model. *Geophys. Res. Lett.* **17**, 1109–1112 (1990).
- Cormier, M.-H. & Macdonald, K. C. East Pacific Rise 18°–19°S: Asymmetric spreading and ridge reorientation by ultrafast migration of axial discontinuities. *J. Geophys. Res.* **99**, 543–564 (1994).
- Cochran, J. R. Variations in subsidence rates along intermediate and fast spreading mid-ocean ridges. *Geophys. J. R. Astr. Soc.* **87**, 421–454 (1986).
- Forsyth, D. W., Dorman, L. M., Phipps Morgan, J. & Shen, Y. Anisotropy and lateral variations in shear wave velocity from propagation of Rayleigh waves across the MELT array. *EOS Transactions AGU* **77**, 652–653 (1996).
- Sandwell, D. T. & Smith, W. H. F. Marine gravity anomaly from Geosat and ERS 1 satellite altimetry. *J. Geophys. Res.* **102**, 10039–10054 (1997).
- Phipps Morgan, J., Parmentier, E. M. & Lin, J. Mechanisms for the origin of mid-ocean ridge axial topography: Implications for the thermal and mechanical structure of accreting plate boundaries. *J. Geophys. Res.* **92**, 12823–12836 (1987).

Acknowledgements. D. Scheirer was helpful in obtaining some of the bathymetry and gravity data used for this study; M. Parmentier assisted us in formulating the conceptual and numerical models. This work was supported by the US NSF.

Correspondence should be addressed to M.A.E. (e-mail: eberle@chimie.univ-nantes.fr).

Faunal turnovers of Palaeogene mammals from the Mongolian Plateau

Jin Meng* & Malcolm C. McKenna†

* Department of Biology and Graduate Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, Massachusetts 01003, USA

† Department of Vertebrate Paleontology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA

Most orders and many families of modern mammals were established during the Palaeogene. Mammalian evolution during this period of time has been correlated with global climatic events^{1–4}, although the timing, mode and scale of such a climate–evolution link remain debatable^{1,5,6}. The Palaeogene global climate was step-punctuated by a warming across the Palaeocene/Eocene boundary about 55 Myr ago and a cool-off throughout the late Eocene and early Oligocene epochs^{1,2,7–10}. The most severe cooling was at 33.5 Myr, slightly after the Eocene/Oligocene boundary, and was characterized by a drop in the mean annual temperature and by changes in vegetation from Eocene dense forests to Oligocene more open country⁵. Here we analyse 33 Palaeogene mammal faunas from the Mongolian Plateau of China and Mongolia. There is a distinct pattern of faunal turnovers: perissodactyl-dominant faunas of the Eocene were abruptly replaced by rodent/lagomorph-dominant faunas of the Oligocene. We interpret the turnovers as having been effected by global climatic shifts and name the prominent biotic reorganization across the Eocene/Oligocene boundary the Mongolian Remodelling, which correlates to the European *Grande Coupure*.

Studies of the effect of climate on mammalian evolution have almost all been made in North America and Europe². To study Palaeogene faunal turnovers in Central Asia and their possible relationships to climatic changes, we analysed 454 species, representing 210 genera and 88 families of mammals, from 33 local faunas in the Mongolian Plateau of China and Mongolia (Fig. 1). Taxonomic names above the species level follow an updated classification of mammals¹¹. We correlated these local faunas using a numerical clustering method (see Methods) and then

converted the clustered faunas to a sequence of local mammal ages on the basis of their stratigraphic occurrences (Fig. 2). We used three radiometric datings and intercontinental faunal correlations to correlate the sequence of relative ages with the geological time-scale and boundaries¹².

In all calculations the late Palaeocene and Oligocene faunas always clustered as one group, whereas the Eocene faunas may or may not form a single group (Fig. 2; Methods), indicating taxonomic differentiations among faunas of the Palaeogene epochs. We concur with placement of the Palaeocene/Eocene boundary (PEB) between the Gashatan and Bumbanian¹³ but note the possibility that it may be above the Bumbanian¹⁴. Our reallocation of the Eocene/Oligocene boundary (EOB) at the Houldjinian/Hsandalgolian transition differs from the traditional placement of the EOB between the Sharamurunian and Ulangochuan^{15,16} but agrees with the ideas of others^{1,17,18}. Our method cannot differentiate Hsandalgolian faunas into the biochronological subunits described in ref. 19, reflecting the faunal stability during that period of time¹⁹. Temporal distributions of important families show that, at the PEB, several archaic taxa (such as multituberculates) became extinct (Fig. 2), whereas Artiodactyla, Euprimates, Condylarthra, true rodents, and Perissodactyla made their first appearance. The Arshantan age witnessed the second significant faunal change, characterized by appearances of several perissodactyl families. The most important biotic reorganization occurred at the EOB: many modern families belonging to rodents, lagomorphs, erinaceid insectivores, ruminant artiodactyls and carnivores first appeared. Future studies may prove the Houldjinian to be only part of the Ergilian¹⁵ and the Kekeamuan a subunit of the Hsandalgolian¹⁸; the former possibility would increase the extinction rate and the latter would decrease the rate of first appearances at the proposed EOB.

The Gashatan and Bumbanian faunas consist mostly of archaic taxa, none of which show dominance (Figs 2 and 3). Despite varying species numbers among faunas, which may be biased by taphonomy and collection or may actually reflect species diversities, all other Eocene faunas are dominated by perissodactyls. The Eocene faunas were then abruptly replaced by Oligocene ones in which rodents and lagomorphs were the prevailing primary consumers, rivalling other taxa not only in species numbers but also in numbers of individual organisms (M.C.M. *et al.*, manuscript in preparation). Except for a single primate species from the Bumbanian fauna²⁰, we have not found any of the euprimates, chiropterans, and plesiadapiformes that are common in the Palaeogene of North America and Europe²¹. The

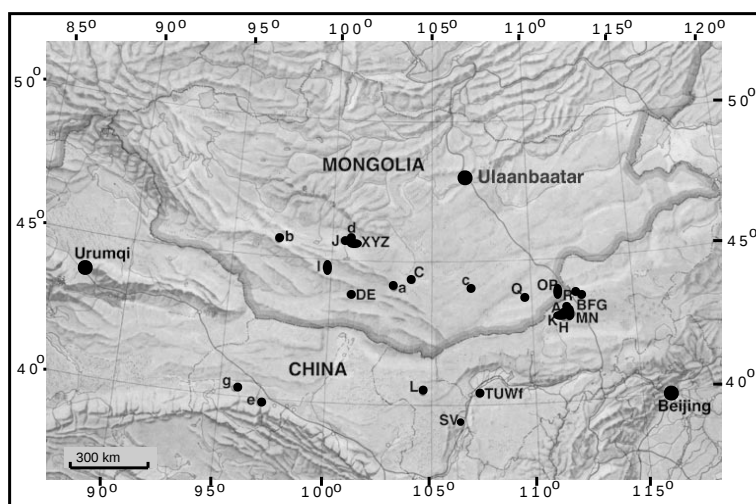


Figure 1 Location map of 33 Palaeogene faunas from the Mongolian Plateau. Faunas are: A, Nomogen; B, Bayan Ulan; C, Gashato; D, Naran; E, Bumban (Tsagan Khushu); F, Arshanto; G, Irdin Manha; H, Ulan Shireh; I, Khaychin; J, Kholboldzi; K, Shara Murun; L, Chaganbulage; M, Ulan Gochu; N, Urtyn Obo; O, Sevkhul at Khoer Dzan; P, Ergilin at Khoer Dzan; Q, Ergilin Dzo; R, Houldjin; S,

Kekeamu; T, Lower Wulanbulage; U, Upper Wulanbulage; V, Ulanatal; W, Saint-Jacques; X, Ulaan Khongil (Tatal Member of Hsanda Gol Formation); Y, Ulaan Khongil (Shand Member); Z, Zavlia (Shand Member); a, Tsakhir; b, Khatan-Khay; c, Shunkht; d, Tsagan Obo; e, Shargaltein; f, Yikebulage; g, Taben Buluk.

apparent lack of these forest-related species implies the absence of dense forests in the region.

Body sizes of species are the most useful single predictor of species adaptation²² and are indicated here by tooth dimensions (Fig. 3; Methods). The Gashatan and Bumban faunas are composed mainly of small and some medium-sized species. The middle and late Eocene faunas show a significant size jump: medium- to large-sized species become common, some of which are among the largest terrestrial mammals ever known. In contrast, the Oligocene faunas experienced a size decrease: large species are few, medium-sized are rare or absent, and small mammals became dominant. The size changes correspond with those of tooth morphologies. As one of the most important groups in the Mongolian Palaeogene, rodents evolved from having a primitive tribosphenic tooth pattern in the late Palaeocene to possessing quadrate-crown molars by the Bumbanian (Fig. 4). Except *Ardynomys* of the family Cylindrodontidae, Eocene rodents have cusps, broadly basined, and low-crowned molars. The Oligocene rodents show a variety of tooth morphologies that are typical of lophodonty, some additionally evolving hypsodonty. Other Oligocene primary consumers, such as ochotonid lagomorphs and ruminant artiodactyls, also developed high-crowned teeth^{19,23}.

Thus, our results reveal a distinctive pattern of faunal turnovers that is reflected by taxonomic differentiation, species appearances and extinctions, reorganization of faunal compositions, changes of species body sizes, and evolution of rodent dentitions. In addition,

major faunal changes occur at short intervals coeval with major climate shifts and are followed by long periods of stability of new communities. Taphonomy, palaeoecology, and methods of collecting are common factors that may bias fossil assemblage data²⁴, as discussed in some studies of Mongolian taxa^{14,15} (M.C.M. *et al.*, manuscript in preparation), but these factors are unlikely to result in such consistent proportions of faunal composition and size structure for the faunas that came from different sediments of the same or different ages; nor does predation account for the onset and crash of the Eocene perissodactyl-dominant faunas. Competition may play only a minor role in shaping the turnover pattern because differences in size and functional morphology between large perissodactyls and small mammals indicate indirect competition for resource use²⁵. Unlike in the European *Grande Coupure*²⁶, the arrival of immigrants as possible competitors is not significant in the Mongolian Palaeogene. Moreover, changes apparently occur across whole faunas, not just in potentially competing lineages.

The Mongolian faunal turnovers are probably effected by major climatic changes. Faunal compositions reflect warm, humid conditions that favoured perissodactyls for most of the Eocene, and a cool, arid environment supporting rodents and lagomorphs during the Oligocene. By comparison with modern and other fossil mammal faunas^{4,27}, the presence of all sizes of species in the Eocene also indicates that Eocene conditions were humid and rich in vegetation, whereas the lack of medium-sized species in the Oligocene indicates that environments were open and arid. In

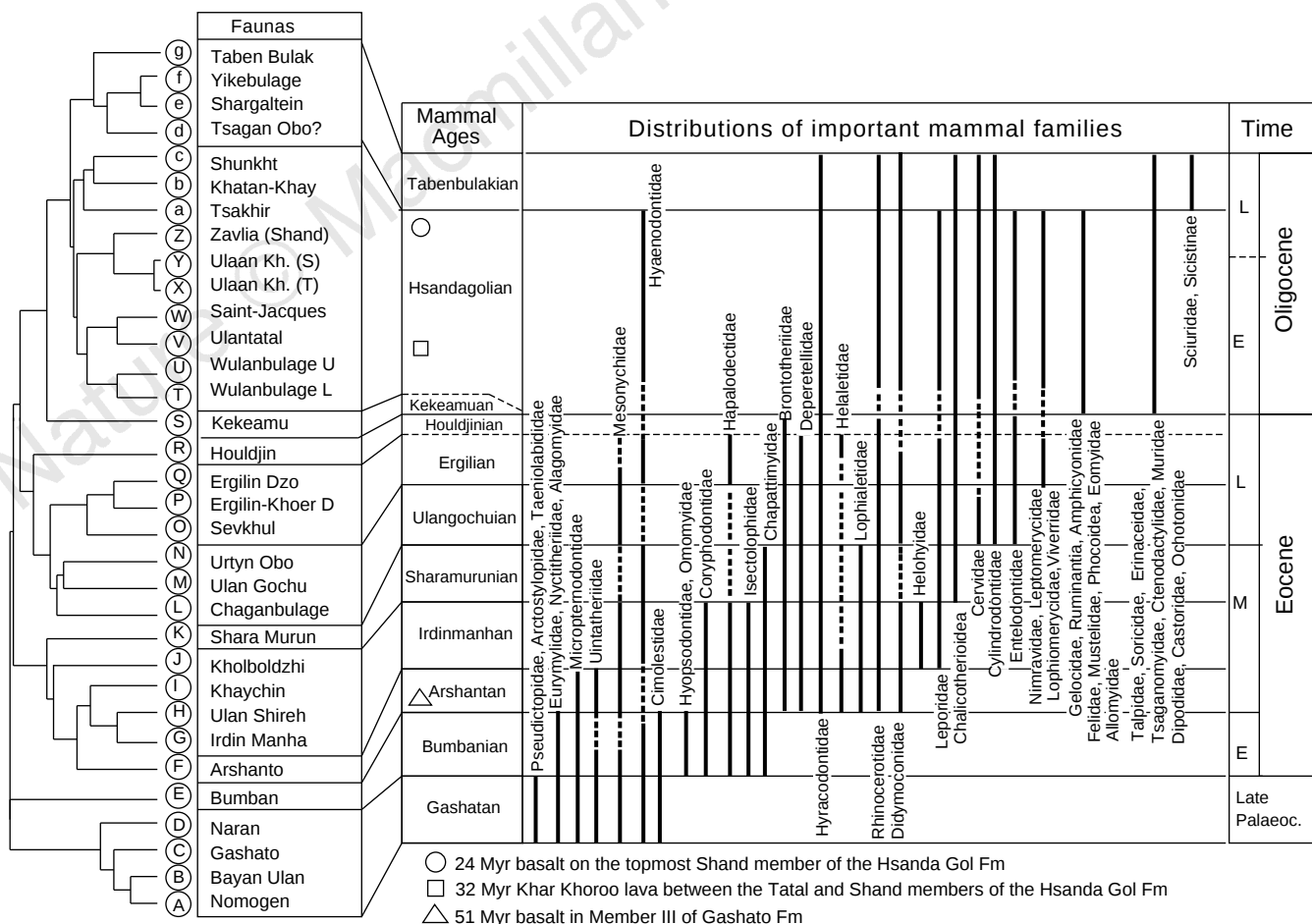


Figure 2 Correlation of local faunas by clustering (see Methods), local mammal ages, and distributions of important mammal families from the Mongolian Palaeogene. Dashed lines indicate no records. The ⁴⁰Ar–³⁹Ar datings of the Khar Khoroo Lava range from 30.9 to 33.7 Myr, whereas the basalt resting at the topmost Shand Member is dated 21–25 Myr (M.C.M. *et al.*, manuscript in

preparation). The less certain 51 ± 3 Myr dating of the basalt in the Gashato Formation is correlated by the presence of *Gomphos elkema* to the top of the Bumban Member in the Naran Bulak Formation¹³. Letters A–Z and a–g at the left relate to locations shown in Fig. 1. Letters at the right are L, late; M, middle; E, early.

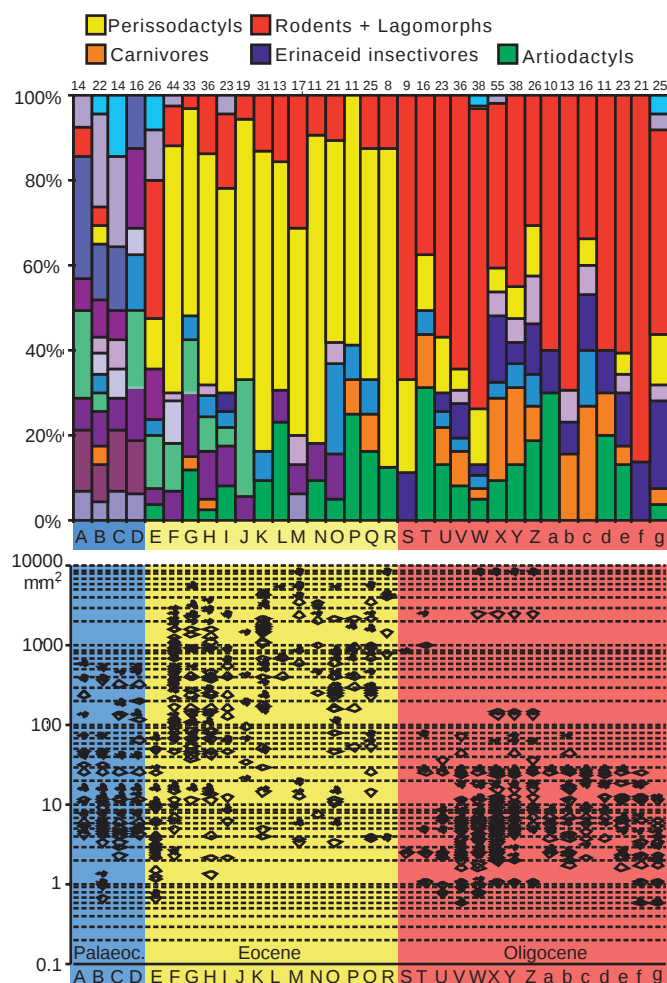


Figure 3 Changes in faunal composition and dental morphology. Top, faunal compositions (in percentages) show the predominance of perissodactyl faunas in the Eocene and of rodent/lagomorph faunas in the Oligocene (complete faunal lists are available as Supplementary Information). Numbers of species for each fauna are at the top of the columns. Bottom, $\log_{10}$ plot of lower and upper molar areas (empty and solid diamonds, respectively; see Methods) in each fauna. We arbitrarily divided the species into three categories: small (TA (upper tooth area) $< 100 \text{ mm}^2$; roughly the size of a small deer or smaller); medium ($100 < \text{TA} < 900 \text{ mm}^2$; approximately up to the size of a tapir); and large ($\text{TA} > 900 \text{ mm}^2$; up to the size of an animal that is 8-m long and 5-m high, about four and a half times as heavy as an adult elephant). There may be biases due to different proportions of tooth-body size between Palaeogene and extant mammals²². Letters on the x-axis correspond to the faunas shown in Figs 1 and 2.

addition, the development of the lophodont and hypsodont teeth in Oligocene rodents and other herbivores probably reflects increased efficiency in grinding abrasive foods²⁸ which are more likely to be expected in arid environments. Therefore, we conclude that the nature of the faunal turnovers of the Mongolian Palaeogene supports the climate/mammalian-evolution link. This pattern is generally comparable with the global framework of mammal evolution, and contrasts with a previous convention²⁹; however, it resembles the European pattern more than the North American one because the most significant turnover in the North American Palaeogene is at the middle-to-late Eocene transition¹. We name the biotic reorganization at the EOB the Mongolian Remodelling and correlate it to the European *Grande Coupure*; both events reflect changes in the environment from a warm, humid Eocene to a cool, arid Oligocene. Because many taxa involved in the Mongolian Remodelling dispersed to Europe at the *Grand Coupure*, the beginning of the

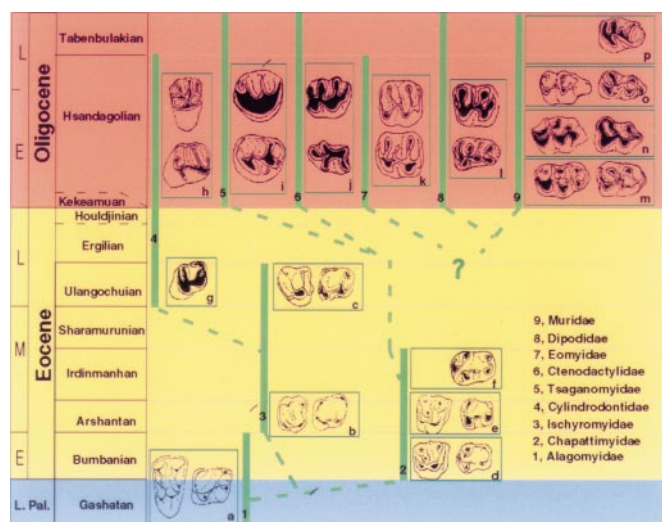


Figure 4 Morphologies of the first right upper and lower molars (anterior to the right; some are photographically reversed; not to scale) of the main Palaeogene rodent families from the Mongolian Plateau. Solid green bars represent distributions and dashed green lines are possible phylogenetic relationships. The species representing each family are: a, *Tribosphenomys minutus*; b, *Asiomys dawsoni*; c, *Hulganina ertnia*; d, *Sharomys singularis*; e, *Tanquammys wilsoni*; f, *Advenimus burkei*; g, *Ardynomys olseni*; h, *Anomoemys lohculus*; i, *Tsaganomys altaicus*; j, *Tataromys plicidens*; k, *Eomys orientalis*; l, *Plesiosminthus tangingoli*; m, *Cricetops dormitor*; n, *Selenomys mimicus*; o, *Eucricetodon asiaticus*; and p, *Tachyoryctoides obrutschewi*. L, late; M, middle; E, early.

Mongolian Remodelling is probably slightly older than the *Grande Coupure*. □

Methods

We clustered faunas using multivariate analysis on an IBM PC³⁰. We used Jaccard and Dice coefficients to build similarity association matrices from binary data (presence and absence of families, genera, and species in faunas). Faunas were then clustered by the unweighted pair-group method (arithmetic average)³⁰. All data generate the main groupings shown in the generic phenogram of Fig. 2, which is the phenogram most congruent with stratigraphical occurrences of faunas. A total of 335 upper and 396 lower molar areas are plotted. Each plotted value is an average of available measurements of molars for a given species. For instance, if only M1 and M2 are available for a species, the upper tooth area is $(M1 \text{ area} + M2 \text{ area})/2$. Each tooth measurement for a species is an average of 1–50 specimens. Most measurements are from literature, supplemented with our own data. Measurements in some taxa, either being studied or fragmentary, are not included; however, expected size distributions of the missing taxa support the plotting results in Fig. 3. Faunal lists and measurements are available as Supplementary Information.

Received 13 November 1997; accepted 11 May 1998.

- Prothero, D. R. & Berggren, W. A. (eds) *Eocene-Oligocene Climatic and Biotic Evolution* (Princeton Univ. Press, Princeton, 1992).
- Janis, C. M. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annu. Rev. Ecol. Syst.* **24**, 467–500 (1993).
- Stucky, R. K. Evolution of land mammal diversity in North America during the Cenozoic. *Curr. Mammal.* **2**, 375–432 (1990).
- Legende, S. & Hartenberger, J.-L. in *Eocene-Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) 516–528 (Princeton Univ. Press, Princeton, 1992).
- Prothero, D. R. & Heaton, T. H. Faunal stability during the early Oligocene climatic crash. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 257–283 (1996).
- Alroy, J. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 285–311 (1996).
- Burchardt, B. Oxygen isotope paleotemperatures from the Tertiary period in the North Sea area. *Nature* **275**, 121–123 (1978).
- Cavelier, C. et al. The geological events at the Eocene/Oligocene boundary. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **36**, 223–248 (1981).
- Shackleton, N. J. et al. Oxygen isotope calibration of the onset of ice-rafting and history of glaciation in the North Atlantic region. *Nature* **307**, 620–623 (1984).
- Rea, D. K. et al. Global change at the Paleocene-Eocene boundary: climatic and evolutionary consequences of tectonic events. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **79**, 117–128 (1990).
- McKenna, M. C. & Bell, S. K. *Classification of Mammals* (Columbia Univ. Press, New York, 1997).

12. Berggren, W. A. *et al.* A revised Cenozoic geochronology and chronostratigraphy. *SEPM Spec. Publ.* **54**, 129–212 (1995).
13. Dashzeveg, D. Holarctic correlation of non-marine Palaeocene-Eocene boundary strata using mammals. *J. Geol. Soc. Lond.* **145**, 473–478 (1988).
14. Beard, K. C. & Dawson, M. R. (eds) Dawn of the Age of Mammals in Asia. *Bull. Carnegie Mus. Nat. Hist.* **34**, 5–39 (1998).
15. Russell, D. E. & Zhai, R. The Palaeogene of Asia: mammals and stratigraphy. *Mem. Mus. Natl. Hist. Nat. C* **52**, 1–488 (1987).
16. Tong, Y., Zheng, S. & Qiu, Z. Cenozoic mammal ages of China. *Vert. Palasiatica* **33**, 290–314 (1995).
17. Ducrocq, S. Mammals and stratigraphy in Asia: is the Eocene-Oligocene boundary at the right place? *C. R. Acad. Sci.* **318**, 549–554 (1993).
18. Wang, B. The mid-Tertiary Ctenodactylidae (Rodentia, Mammalia) of eastern and central Asia. *Bull. Am. Mus. Nat. Hist.* **234**, 1–88 (1997).
19. Vislobokova, I. A. Age of the Shand Gol fauna and evolution of central Asia mammals in the Oligocene. *Stratigr. Geol. Correl.* **4**, 156–165 (1996).
20. Dashzeveg, D. & McKenna, M. C. Tarsioid primate from the early Tertiary of the Mongolian People's Republic. *Acta Palaeontol. Pol.* **22**, 119–137 (1977).
21. Savage, D. E. & Russell, D. E. *Mammalian Paleofaunas of the World* (Addison-Wesley, London, 1983).
22. Damuth, J. & MacFadden, B. J. (eds) *Body Size in Mammalian Paleobiology* (Cambridge Univ. Press, New York, 1990).
23. Jernvall, J., Hunter, J. P. & Fortelius, M. Molar tooth diversity, disparity, and ecology in Cenozoic ungulate radiations. *Science* **274**, 1489–1492 (1996).
24. Badgley, C. *et al.* Taphonomy of vertebrate assemblages from the Paleogene of northwest Wyoming and the Neogene of northern Pakistan. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **115**, 157–180 (1995).
25. Morgan, M. E. *et al.* Comparative paleoecology of Paleocene and Neogene mammalian faunas: body-size structure. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **115**, 287–317 (1995).
26. Stehlin, H. G. Remarques sur les faunules de mammifères des couches eocènes et oligocènes du Bassin de Paris. *Bull. Soc. Geol. Fr.* **9**, 488–520 (1909).
27. Legendre, S. Analysis of mammalian communities from the late Eocene and Oligocene of southern France. *Palaeovertebrata* **16**, 191–212 (1986).
28. Butler, P. M. in *Evolutionary Relationships among Rodents* (eds Luckett, W. P. & Hartenberger, J.-L.) 381–401 (Plenum, New York, 1985).
29. Dashzeveg, D. Asynchronism of the main mammalian faunal events near the Eocene-Oligocene boundary. *Tertiary Res.* **14**, 141–149 (1993).
30. Rohlf, F. J. *NTSYS-pc Numerical Taxonomy and Multivariate Analysis System* (Applied Biostatistics, New York, 1994).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank B. Wang for unpublished data; R. H. Tedford, X. Wang, M. Coombs, G. Gould, K. Doyle and the OEB discussion group of the University of Massachusetts for discussions; and J.-L. Hartenberger for comments. This work was supported by the NSF, the National Geographic Society, the University of Massachusetts and the American Museum of Natural History.

Correspondence and requests for materials should be addressed to J.M. (e-mail: jmeng@bio.umass.edu). Supplementary data are available at: <http://www.nature.com>.

Veil architecture in a sulphide-oxidizing bacterium enhances countercurrent flux

Tom Fenchel & Ronnie N. Glud

Marine Biological Laboratory, University of Copenhagen, 3000 Helsingør, Denmark

Solute uptake by microorganisms is limited by molecular diffusion through a boundary layer surrounding the cells, and the uptake is not enhanced (or only insignificantly) by convective water transport or by swimming¹. It is generally assumed that sediment uptake of oxygen is diffusion-limited, so the steepness of the concentration gradient within the 0.5–1-mm-thick diffusive boundary layer is a measure of diffusional flux into the sediment^{2,3}. Here we show that veils, which are formed on sediments by the marine sulphide-oxidizing bacterium *Thiovulum majus*⁴, generate convective oxygen transport through the 0.5-mm-thick water layers above the veil at rates that are about 40 times higher than molecular diffusion. Chemosensory behaviour of the cells, combined with their generation of water currents, leads to characteristic, aggregated distribution patterns: areas with high cell densities draw oxygenated water downwards through the veil, whereas areas without cells serve for the upward-directed return flow of deoxygenated water. The microbial community structure thus overcomes the limitations of diffusion and thereby enhances the rates of respiration and sulphide oxidation.

Thiovulum majus forms characteristic white veils on or slightly above sulphidic sediments. The cells are unusually large (5–10 µm diameter) and they are among the fastest-swimming bacteria

known (150–600 µm s⁻¹)⁵. Their motility has both a translational and a rotational component resulting in a helicoid swimming path. The cells show a pronounced chemosensory behaviour in O₂ gradients, in which they form a narrow band at an O₂ tension of about 4% atmospheric saturation (0.85 kPa); in nature this ensures that the bacteria aggregate in a <100-µm-thick zone where oxygen and sulphide coexist^{4–7}.

An approximately 100-µm-thick band of swimming *Thiovulum* cells, maintained by a steep O₂ gradient within a glass capillary, caused no detectable displacement of suspended latex particles outside the band. Within the band, individual latex particles were displaced during close encounters with swimming cells. Among 20 such encounters, the maximum displacement (λ) recorded was 10 µm, and the duration of the displacements (τ) was 0.3–0.5 s. A crude estimate of diffusive mixing due to bacterial swimming is then given by $\lambda^2/4\tau$, or about 6×10^{-7} cm² s⁻¹. This value, which represents an overestimate because only a fraction of the beads were moving at any time, is much lower than diffusivity of dissolved O₂. These results accord with theoretical hydrodynamics: non-thrust swimming at low Reynolds numbers causes only a local flow field (less than the length of the translating cell), whereas tethered cells cause a large-scale flow field⁸. In addition, swimming cells should not cause a large-scale convective flow because they swim in random directions within the band⁶.

Under optimal conditions, cells secrete a mucous thread from their posterior end. The threads, which may exceed 100 µm in length, tend to attach to solid surfaces, but the cells detach themselves from the thread if the O₂ tension deviates from the optimal value⁶. Attached cells are typically oriented towards the oxic side of the chemical gradients. The cells maintain their rotational motility (with an angular velocity of ~ 30 s⁻¹), but translational motion is prevented by the stretched thread. Instead, the cells create a flow field with a velocity of ~ 200 µm s⁻¹ close to the cell (Fig. 1a).

The attached bacteria form characteristic white veils on natural sulphidic sediments. These are sometimes constituted by equally spaced cells punctuated by holes measuring 100–500 µm in diameter and with a raised rim (Fig. 2). Alternatively, the cells occur in small, regularly spaced clusters, each containing 25–100 cells (Fig. 3). The type of veil that develops is a function of the height of the O₂–H₂S transition zone above the sediment: continuous veils with more or less regular holes occur mainly over depressions in the sediment surface, whereas isolated cell clusters appear on slightly elevated areas.

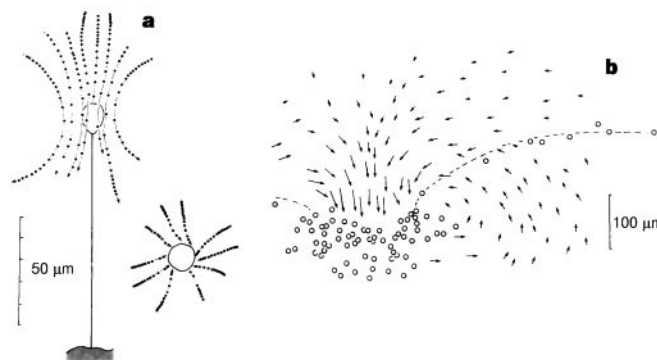


Figure 1 The flow field around *Thiovulum* cells. **a**, Flow field around two isolated and attached cells seen from the side (left) and from above (right), respectively. Time interval between positions of latex beads (black circles) is 0.04 s. **b**, Side view of the flow field (velocity vectors) around a group of attached cells; the posterior and anterior ends of the arrows indicate the positions of a bead before and after an interval of 0.2 s. The dashed line reflects the horizon where swimming *Thiovulum* cells occurred; it corresponds to an O₂ tension of approximately 4% atmospheric saturation.

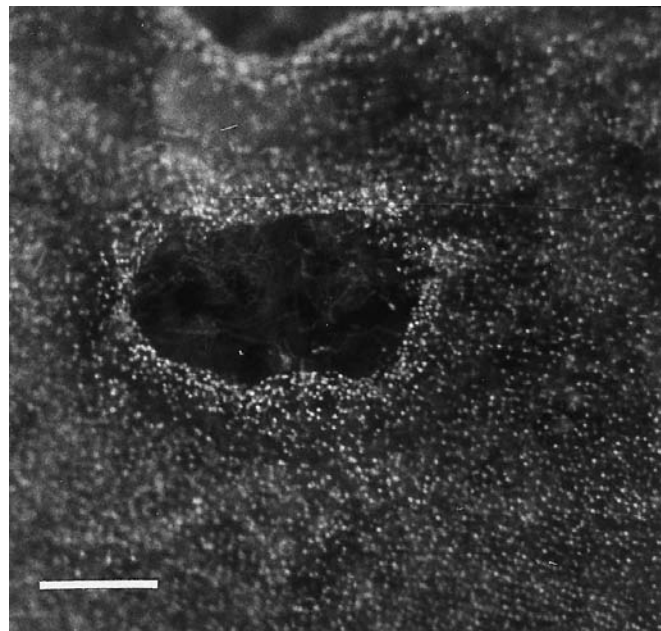


Figure 2 A continuous *Thiovulum* veil with holes. Scale bar, 0.25 mm.

The flow field around a cell cluster close to a solid surface is characterized by a downward current of oxic water which passes into the cell cluster and out in a horizontal and upwards direction (Fig. 1b). Free-swimming cells occur around the attached clusters. The swimming cells follow the 4% atmospheric saturation O_2 isopleth. The horizon with free-swimming cells moves upwards outside the clusters of attached cells; this indicates that deoxygenated water flows upwards outside the clusters (Figs 1b and 2). In continuous veils with holes, water is drawn downwards and through the veil, whereas the holes serve as outflows for deoxygenated water. Consequently, the oxygen isopleths show a dome structure over the holes and are depressed over the bacterial veil (Fig. 4). Along the rim of the holes, the veils bend upwards, following the optimum O_2 tension. Above the veil, the upward flow of deoxygenated water and the downward flowing oxygenated water partly mix; this, rather than diffusive flux, is responsible for the O_2 concentration gradient above the veil. The upward flowing water is probably also slightly sulphidic, and so the bacteria also receive some sulphide in addition to oxygen through the incoming flow.

The patterns formed by the bacteria are easily understood in terms of their chemosensory behaviour. A discrete cell cluster with short attachment threads disposes of deoxygenated water laterally and this prevents newcoming cells from attaching in the immediate vicinity of the established group. If the sulphidic layer extends somewhat higher above the sediment surface, the attachment threads are longer. The incoming waterflow therefore has an almost vertical direction at the level of the cells, and this allows for a more continuous veil. There must be a compensatory outflow, however, and so holes form. These are kept open because anoxia extends up into the water column above the level of the veil.

It is possible to estimate the magnitude of convective transport relative to diffusive transport. The Sherwood number (kL/D , where k is convective transport, or vertical flow velocity within the veil, L is distance and D is the diffusion coefficient) is a dimensionless measure of the relative role of convective and diffusive transport⁹. If $k = 0.015 \text{ cm s}^{-1}$, $D = 2 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$ (at $\sim 20^\circ \text{C}$) and $L = 0.05 \text{ cm}$ (the distance above the bacterial veil at which turbulent mixing almost maintains atmospheric O_2 saturation), this works out to ~ 40 , indicating that convective transport is more important than diffusion for supplying O_2 to the bacteria. The result of the convective transport is an enhanced rate of sulphide oxidation and

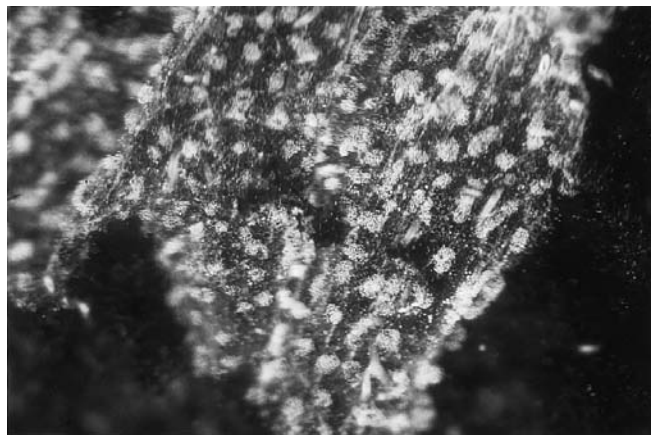


Figure 3 Regularly spaced groups of *Thiovulum* cells on a detrital particle (dead eelgrass leaf tissue).

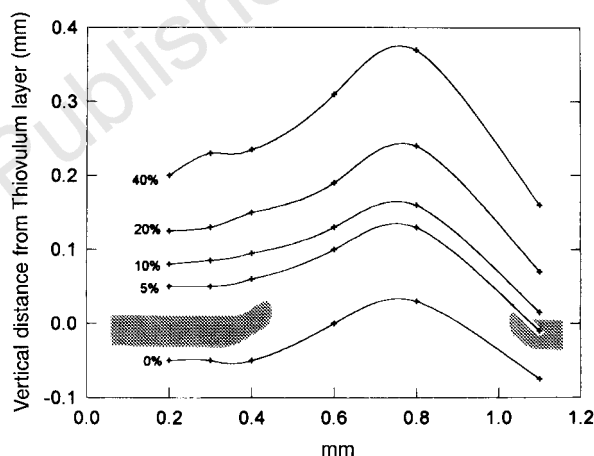


Figure 4 Oxygen isopleths (expressed as per cent of atmospheric saturation) above a hole in a *Thiovulum* veil; shading indicates the position of the veil.

oxygen uptake by sediment, and it allows for a higher density of bacteria. By far the greatest fraction of O_2 uptake in such veils is due to sulphide oxidation by the chemolithotrophic sulphur bacteria¹⁰.

Convective oxygen transport in the surface layer of sediments has been shown to result from bioturbation, methane ebullition, or seepage of hot water in hydrothermal vents; it is also recognized that this may enhance microbial activity^{11–13}. To our knowledge, we have provided the first example of bacteria developing spatially organized structures in order to generate convective solute transport. □

Methods

Material. Samples of sulphidic sediment collected in jars from the innermost part of Nivå Bay ($\sim 25 \text{ km}$ north of Copenhagen) remained a reliable source of *Thiovulum* for several weeks⁶. Freshly collected sediment cores with white patches of *Thiovulum* were used for photographic evidence of veil structure and for measuring O_2 gradients. The supernatant water was bubbled with air to prevent the bacteria from migrating up in the water column in response to anoxia.

Flow fields. Bacteria were allowed to form veils, either in ordinary microscopic preparations (with a distance between slide and coverslip of $0.5\text{--}1 \text{ mm}$) or in flat, capillary 'microslides' (Camlab, Cambridge); the internal dimensions were $4 \text{ mm} \times 0.4 \text{ mm}$. To study free-swimming cells, we filled the capillary with bacteria in a suspension of latex beads ($1 \mu\text{m}$); when the cells had used most of the dissolved O_2 , they formed a narrow band at $0.4\text{--}0.9 \text{ mm}$ from the meniscus (*Thiovulum* rarely attaches to glass surfaces). Attached cells were studied by the addition of small pieces of detritus to which the bacteria could attach. The bacteria consume the oxygen in the preparation and thus create steep O_2

gradients close to the air–water interface, where veils develop within a few minutes. A suspension of latex beads (1 or 0.5 μm) was then added with a capillary pipette. Preparations were recorded with dark-field illumination using a CCD camera (Sony) attached to the microscope and a video recorder. Afterwards, positions of individual latex beads were recorded frame by frame (at 0.04-s intervals). Only trajectories that were almost parallel to the plane of observation were included.

Oxygen gradients. Oxygen microelectrodes (with a 5–10- μm tip)³ connected to a picoammeter were mounted in a micromanipulator. The surface of the *Thiovulum* veil and the electrode tip were observed through a dissection microscope tilted at an angle of 45°; this made it possible to map the O_2 isopleths inside and above the veil. Microelectrodes penetrating from above tend to deform the O_2 diffusion gradients above the sediment¹⁴, but this should not affect the relative position of O_2 isopleths in different areas of the veil. The fact that O_2 was measured in a vertical convective flow is also likely to minimize this effect.

Received 27 January; accepted 8 May 1998.

1. Purcell, E. M. Life at low Reynolds number. *Am. J. Phys.* **45**, 3–11 (1977).
2. Revsbech, N. P., Jørgensen, B. B. & Blackburn, T. H. Oxygen in the sea bottom measured with a microelectrode. *Science* **207**, 1355–1356 (1980).
3. Revsbech, N. P. & Jørgensen, B. B. Microelectrodes: their use in microbial ecology. *Adv. Microbiol. Ecol.* **9**, 293–352 (1986).
4. Wirsén, C. O. & Jannasch, H. W. Physiological and morphological observations on *Thiovulum* sp. *J. Bacteriol.* **136**, 765–774 (1978).
5. Garcia-Pichel, F. Rapid bacterial swimming measured in swarming cells of *Thiovulum majus*. *J. Bacteriol.* **171**, 3560–3563 (1989).
6. Fenchel, T. Motility and chemosensory behaviour of the sulphur bacterium *Thiovulum majus*. *Microbiology* **140**, 3109–3116 (1994).
7. Jørgensen, B. B. & Revsbech, N. P. Colorless sulfur bacteria, *Beggiatoa* spp. and *Thiovulum* spp. in O_2 and H_2S microgradients. *Appl. Environ. Microbiol.* **45**, 1261–1270 (1983).
8. Lighthill, J. Flagellar hydrodynamics. *Soc. Indust. Appl. Math. Rev.* **18**, 161–230 (1976).
9. Vogel, S. *Life in moving fluids* (Grant, Boston, 1981).
10. Fenchel, T. & Bernard, C. Mats of colourless sulphur bacteria. I Major microbial processes. *Mar. Ecol. Prog. Ser.* **128**, 161–170 (1995).
11. Aller, R. C. & Yingst, J. Y. Effects of marine deposit-feeders *Heteromastus filiformis* (Polychaeta), *Macoma balthica* (Bivalvia), and *Tellina texana* (Bivalvia) on averaged solute transport, reaction rates, and microbial distributions. *J. Mar. Res.* **41**, 299–322 (1985).
12. Dando, P. R. *et al.* The effects of methane seepage at an intertidal/shallow subtidal site on the shore of Kattegat, Vendsyssel. *Bull. Geol. Soc. Denmark* **41**, 65–79 (1994).
13. Gundersen, J. K., Jørgensen, B. B., Larsen, E. & Jannasch, H. W. Mats of giant sulphur bacteria on deep-sea sediments due to fluctuating hydrothermal flow. *Nature* **360**, 454–455 (1992).
14. Glud, R. N., Gundersen, J. K., Revsbech, N. P. & Jørgensen, B. B. Effects on the benthic diffusive boundary layer imposed by microelectrodes. *Limnol. Oceanogr.* **39**, 462–467 (1994).

Acknowledgements. We thank J. Johansen for technical assistance. This study was supported by grants from the Danish Natural Science Research Council.

Correspondence and requests for materials should be addressed to T.F. (e-mail: mbltf@inet.uni2.dk).

Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei

T. Wakayama*†, A. C. F. Perry*‡, M. Zuccotti*§, K. R. Johnson|| & R. Yanagimachi*

* Department of Anatomy and Reproductive Biology, John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii 96822, USA

† Department of Veterinary Anatomy, Faculty of Agriculture, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

‡ Department of Signalling, Babraham Institute, Cambridge CB2 4AT, UK

§ Dipartimento Biologia Animale, Laboratorio Biologia dello Sviluppo,

University of Pavia, Piazza Botta 10, 27100, Pavia, Italy

|| Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609, USA

Until recently, fertilization was the only way to produce viable mammalian offspring, a process implicitly involving male and female gametes. However, techniques involving fusion of embryonic or fetal somatic cells with enucleated oocytes have become steadily more successful in generating cloned young^{1–3}. Dolly the sheep⁴ was produced by electrofusion of sheep mammary-derived

cells with enucleated sheep oocytes. Here we investigate the factors governing embryonic development by introducing nuclei from somatic cells (Sertoli, neuronal and cumulus cells) taken from adult mice into enucleated mouse oocytes. We found that some enucleated oocytes receiving Sertoli or neuronal nuclei developed *in vitro* and implanted following transfer, but none developed beyond 8.5 days post coitum; however, a high percentage of enucleated oocytes receiving cumulus nuclei developed *in vitro*. Once transferred, many of these embryos implanted and, although most were subsequently resorbed, a significant proportion (2 to 2.8%) developed to term. These experiments show that for mammals, nuclei from terminally differentiated, adult somatic cells of known phenotype introduced into enucleated oocytes are capable of supporting full development.

Previous studies have suggested that embryonic development is enhanced when donor nuclei are in the G0 or G1 phase of the cell cycle^{1,3,4}, and Dolly the sheep developed from an enucleated oocyte electrofused with a mammary-derived cell presumed to be in G0 following culture in serum-deficient medium for 5 days⁴. We have investigated the developmental potential of oocytes injected with the nuclei of non-cultured cells known to be at G0. We selected

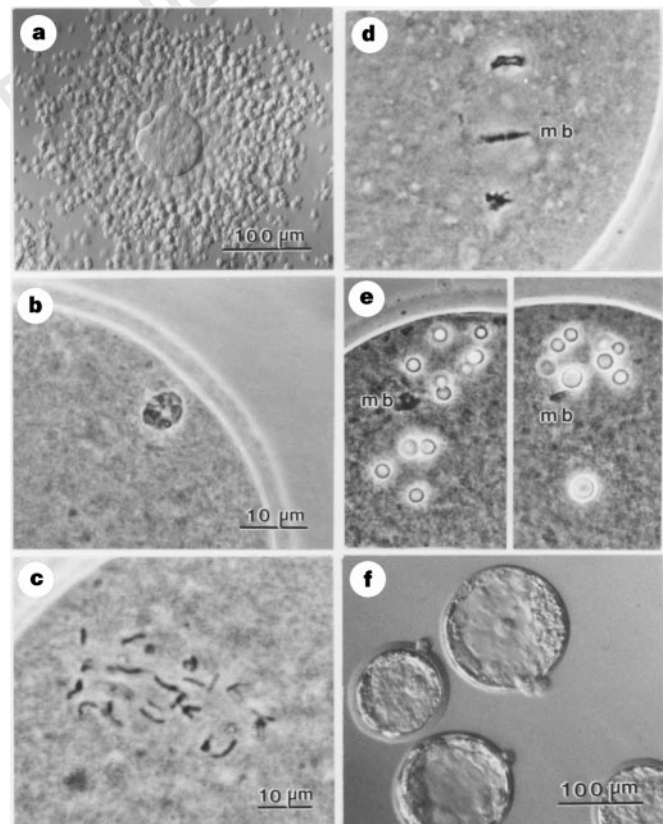


Figure 1 *In vitro* development of enucleated oocytes following injection of cumulus cell nuclei. **a**, Live oocyte surrounded by cumulus cells. The egg coat (the zona pellucida) appears in this micrograph as a relatively clear zone around the oocyte. **b–e**, Behaviour of cumulus cell nuclei following injection into enucleated oocytes, photographed after fixation and staining. **b**, A cumulus cell nucleus within 10 min of injection. **c**, Transformation of the nucleus into disarrayed chromosomes 3 h after injection. The disorder reflects an unusual situation in which single, condensed chromatids are each attached to a single pole of the spindle and are therefore not aligned on a metaphase plate. **d**, 1 h after Sr^{2+} activation, chromosomes are segregated into two groups (mb, midbody). **e**, 5 h after Sr^{2+} activation, two pseudo-pronuclei (left and right panels) with a varying number of distinct nucleolus-like structures are discernible in each egg. The size and number of pseudo-pronuclei varied, suggesting that segregation of chromosomes was random after oocyte activation. **f**, Live blastocysts produced following injection of enucleated oocytes with cumulus cell nuclei.

Table 1 Preimplantation of enucleated eggs injected with cumulus cell nuclei

Time of oocyte activation	Total no. of oocytes used	No. of enucleated oocytes	No. of surviving oocytes after injection	No. (%) of activated oocytes	No. (% mean $\pm$ s.d.) of embryos developed from oocytes at 72 h after activation		
					1-cell and abnormal	2-8-cell	Morula/blastocyst*
Simultaneously with injection	233	230	182	153 (84.1)	17	75	61 (39.9 $\pm$ 16.6)
1-3 h after injection	573	565	508	474 (93.3)	20	177	277 (58.4 $\pm$ 12.6)
3-6 h after injection	195	191	182	151 (83.0)	9	41	101 (66.9 $\pm$ 14.4)

* There is a significant different ($P < 0.005$) between the top result and the bottom two. Data were analysed using the χ^2 test.

Sertoli, neuronal and cumulus from adult mice as representatives of this class; Sertoli cells and neurons do not normally divide in adults but remain at G0, and more than 90% of cumulus cells surrounding recently ovulated oocytes (Fig. 1a) are in the G0/G1 phase of the cell cycle⁵. These somatic cell types have very distinctive morphologies, making them easy to identify with confidence. All cells were used immediately (that is, without *in vitro* culturing) following the removal of tissue from freshly killed mice.

Enucleated mouse oocytes were each injected with a single

nucleus from one of the three somatic cell types and left for 0 to 6 hours before activation. Examination of enucleated oocytes injected with cumulus nuclei revealed that chromosome condensation had occurred within 1 hour of injection (Fig. 1b, c). When, after 1 to 6 hours incubation, oocytes were activated in culture medium containing Sr^{2+} and cytochalasin B, their cumulus-derived chromosomes segregated (Fig. 1d) to form structures resembling the pronuclei that are formed after normal fertilization (referred to here as pseudo-pronuclei). Examination of 47 such oocytes after fixation and staining showed that 64% had two pseudo-pronuclei (Fig. 1e) and 36% had three or more. Oocytes with distinct pseudo-pronuclei were considered to be activated. Owing to the cytokinesis-blocking effect of cytochalasin B, no polar body was formed and therefore all chromosomes were retained within the oocyte, regardless of the number of pseudo-pronuclei. Chromosome analysis of 13 such oocytes fixed before the first cleavage (data not shown) revealed that 85% had a normal total chromosome number ($2n = 40$). The time interval between nucleus injection and oocyte activation appeared to affect the rate of oocyte development (Table 1). Activation immediately after nucleus injection led to

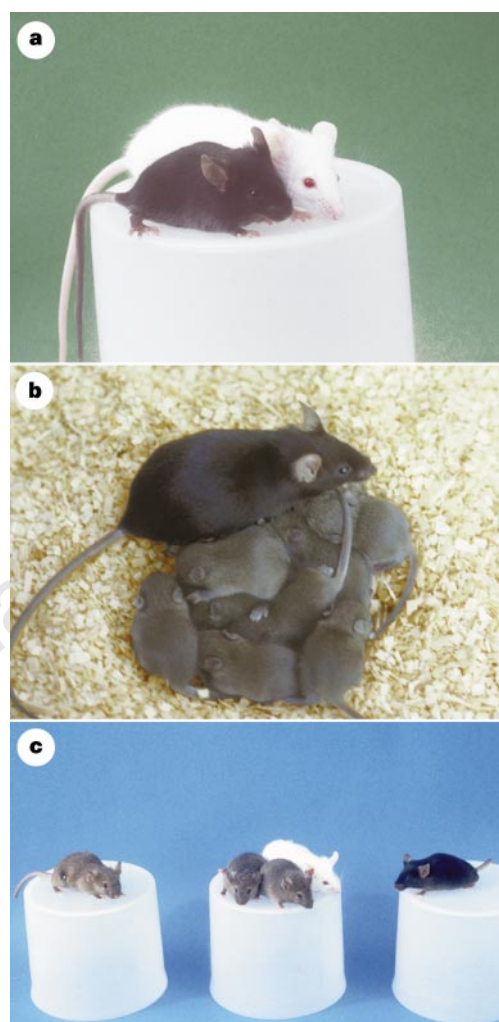


Figure 2 Cloned mice. **a**, The first surviving cloned mouse, Cumulina (born 3 October 1997) at four weeks (foreground) with her foster mother. **b**, Cumulina at 2.5 months with the pups she produced following mating with a CD-1 (albino) male. **c**, Two B6C3F1-derived, cloned, agouti young (centre) in front of their albino foster mother (CD-1), and a B6D2F1 oocyte donor (black, right). The two agouti offspring in the centre are clones (identical 'twin' sisters) of the agouti B6C3F1 cumulus donor shown on the left, and are two of the offspring described in series C (see text) and Table 2.

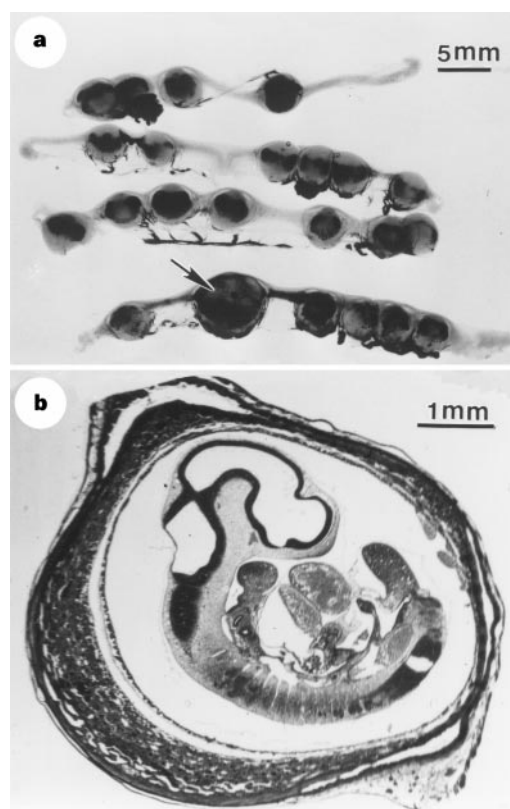


Figure 3 Development following uterine transfer of embryos produced after injection of Sertoli cell nuclei into enucleated oocytes. **a**, Uteri of recipient females 8.5 d.p.c., fixed with Bouin's fluid, dehydrated and cleared with benzyl benzoate. All uterine implantation sites failed to develop except for one (arrow), in which an embryo (**b**) appeared to be normal and was at the ~12-somite stage.

Table 2 Development of enucleated eggs injected with Sertoli or brain-cell nuclei*

Cell type injected	No. of surviving oocytes injected	No. (%) of oocytes activated	Total no. (%) of morulae/blastocysts developed†	No. of transferred embryos (recipients)	No. (%) of implantation sites	No. (%) of fetuses
Sertoli	159	159 (100)	63 (39.6)	59 (8)	41 (69.5)	1 (1.7)
Brain	228	223 (97.8)	50 (22.4)	46 (5)	25 (54.3)	1 (2.2)‡

* All recipients were killed at 8.5 d.p.c.

† There is a significant difference ($P < 0.005$) between the top and bottom result.

‡ Died at about 6–7 d.p.c.

significantly poorer development to morulae/blastocysts *in vitro* (Fig. 1f) than was achieved when activation followed a delay of 1 to 6 hours.

On the basis of this information, we injected Sertoli and neuronal nuclei into enucleated oocytes and delayed activation for 1 to 6 hours: about 40% of enucleated oocytes that had been injected with Sertoli cell nuclei, and 22% of those injected with neuronal nuclei, developed into morulae/blastocysts *in vitro* (Table 2). As these values were less than those achieved after cumulus nucleus injection (58–67%; Table 1), we concentrated on the potential of cumulus cell nuclei to support embryonic development *in vivo*.

In the first series of experiments (series A in Table 3), a total of 142 developing embryos (at the 2-cell to blastocyst stage) were transferred to 16 recipient females. When these females were examined at 8.5 and 11.5 d.p.c., 5 live and 5 dead fetuses were seen *in utero*. In the second series (series B in Table 3), a total of 800 embryos were transferred into 54 foster mothers, and caesarean sections at 18.5–19.5 d.p.c. revealed 17 live fetuses. Of these, six died soon after delivery, one died approximately 7 days after delivery, but the remaining ten females survived and are apparently healthy. All of these, including the first-born survivor (born on 3 October 1997 and named 'Cumulina'; Fig. 2a), have been mated and have delivered and raised normal offspring (Fig. 2b). Several of these offspring have, in turn, now developed into fertile adults.

In the third series of experiments (series C in Table 3), B6C3F1 mice carry a copy of the *agouti* (A) gene, and are consequently agouti; offspring from this experiment should therefore have an agouti coat colour, rather than the black of the B6D2F1 oocyte donors. A total of 298 embryos derived from B6C3F1 cumulus cell nuclei were transferred to 18 foster mothers. Caesarean sections at

19.5 d.p.c. revealed six live fetuses whose placentas were used in DNA-typing analysis. Although one died a day after birth, the five extant females are healthy and have the agouti coat phenotype. Figure 2c shows two such agouti pups with their albino foster mother (CD-1).

We did additional experiments (series D in Table 3) to investigate whether clones could be more efficiently cloned in subsequent rounds of recloning. We therefore collected cumulus cells from B6C3F1 (agouti) clones generated in series C and injected their nuclei into enucleated B6D2F1 oocytes to generate embryos that were transferred as described for series B and C. A total of 287 embryos derived from cloned B6C3F1 cumulus-cell nuclei were transferred to 18 foster mothers. When caesarean sections were done at 19.5 d.p.c. eight live fetuses were recovered. Although one died soon after birth, the seven surviving females are healthy and have the predicted agouti coat phenotype. These results indicate that clones (series B and C) and cloned clones (series D) are produced with comparable efficiency. This argues that successive generations of clones do not undergo changes (either positive or negative) that influence the outcome of the cloning process.

We also monitored the developmental potential *in vivo* of morulae/blastocysts generated following the injection of either Sertoli-cell or neuronal nuclei into enucleated oocytes (all cells from non-clones) (Table 2). Embryos produced by Sertoli-nucleus injection resulted in a single live fetus (Fig. 3) in the uterus of a foster mother killed 8.5 d.p.c. (Table 2). We failed to detect *in vivo* development of embryos derived following injection of neuronal nuclei beyond 6–7 d.p.c.

We believe that all of the live offspring reported here represent clones derived from cumulus-cell nuclei in the absence of genetic

Table 3 Postimplantation development of enucleated eggs injected with cumulus cell nuclei

Experiment series*	Time of oocyte activation	No. of injected oocytes	No. of transferred embryos (recipients)	No. (%) of implantations from transferred embryos†	No. of fetuses developed from transferred embryos					No. (%) of newborn from transferred embryos
					Total (%)†	8.5 d.p.c.		11.5 d.p.c.		
						Live	Dead	Live	Dead	
A	Simultaneously with injection	82	34 (4)	8 (23.5)	0					–
	1–3 h after injection	136	45 (5)	32 (71.1)	7 (15.6)	3	2‡	2	0	–
	3–6 h after injection	124	63 (7)	36 (57.1)	3 (4.8)	0	2§	0	1	–
B	1–3 h after injection	1345	760 (49)	–	–	–	–	–	–	16 (2.1)
	3–6 h after injection	62	40 (5)	–	–	–	–	–	–	1 (2.5)
C	1–3 h after injection	458	298 (18)	–	–	–	–	–	–	6 (2.0)
D	1–3 h after injection	603	287 (18)	–	–	–	–	–	–	8 (2.8)

* Series A, caesarean sections were done at 8.5 or 11.5 d.p.c.; series B and C, caesarean sections were done at 18.5–19.5 d.p.c. In series A and B, each donor nucleus is from a B6D2F1 cumulus cells; in series C, each donor nucleus is from a B6C3F1 cumulus cell; in series D, each donor nucleus is from a B6C3F1 clones mouse from series C.

† There is a significant difference between the top result and the bottom two: implantation ($P < 0.005$); fetal development ($P < 0.05$). Data were analysed by χ^2 tests.

‡ Died 6–7 d.p.c.

§ Died 7–8 d.p.c.

|| Died 10 d.p.c.

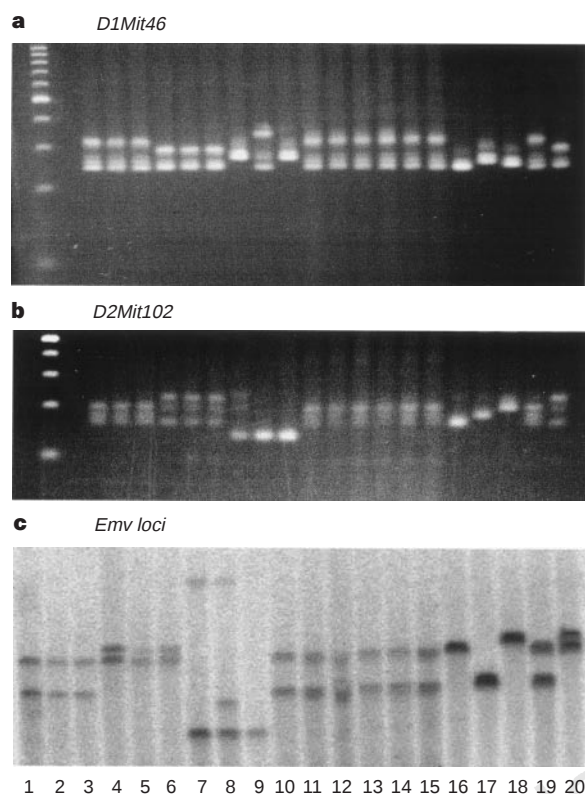


Figure 4 DNA typing of donors and offspring in series C corroborates the genetic identity of the cloned offspring to cumulus cell donors, and non-identity to oocyte donors and host foster females. **a**, PCR typing using the strain-specific marker *D1Mit46*. **b**, PCR-amplified DNA (**a**, **b**) from F_1 hybrid mice gives an additional gel band not seen in the DNA from inbred parental strains (lanes 16–20); this extra band corresponds to a heteroduplex derived from the two parental products, whose conformation results in anomalous gel migration. **c**, Southern blot typing of strain-specific *Emv* loci (*Emv1*, *Emv2* and *Emv3*). Placental DNA from the six cloned series C offspring (lanes 10–15) was compared with DNA from the three cumulus cell donor females (lanes 1–3), the three oocyte recipient females (lanes 4–6), and the three host females (lanes 7–9). Control DNA was from C57BL/6 (lane 16), C3H (lane 17), DBA/2 (lane 18), B6C3F1 (lane 19) or B6D2F1 (lane 20). 100-bp DNA size-marker ladders are shown on the left of **a** and **b**.

contamination, for the following reasons. (1) Oocytes/eggs were not exposed to spermatozoa *in vitro*. (2) Foster mothers (CD-1, albino) were mated with vasectomized males (CD-1, albino) of established infertility. In the unlikely event of fertilization by such a vasectomized male, the offspring would be albino. We transferred 2- to 8-cell embryos/blastocysts into the oviducts/uteri of foster mothers; it is well established that 2- to 8-cell mouse embryos/blastocysts cannot be fertilized by spermatozoa⁶. (3) All full-term animals were born with black eyes; the surviving ten from series B have black coats and the surviving five in series C have agouti coats. This pattern of coat colour inheritance exactly matches that predicted by the genotype of the nucleus donor in each case. As B6D2F1 mice lack the *agouti* gene, the agouti mice in series C must have inherited their agouti coat colour from a non-B6D2F1 nucleus. (4) Where possible, all putative clones have been sexed, and all were found to be females, consistent with their genetic progenitors invariably being female. (5) DNA typing of highly variable alleles diagnostic of the B6, C3, D2 and CD-1 strains used here (Fig. 4) demonstrates beyond reasonable doubt that the six cloned offspring in series C (which includes one that died soon after birth) are isogenic with the three cumulus cell donor females used (B6C3F1) and do not contain DNA derived from either the oocyte donors (B6D2F1) or host foster mothers (CD-1). (6) Following enucleation, we suppressed extru-

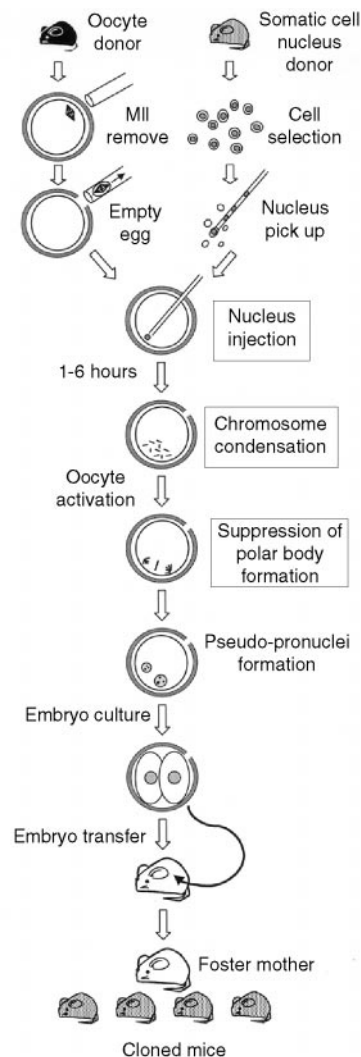


Figure 5 The cloning procedure developed here, as described in the text and Methods section.

sion of chromosomes into polar bodies using cytochalasin B. Thus, even if enucleation of the oocytes had been either totally unsuccessful or only partly successful, all resulting zygotes would be hyperploid; such embryos are inviable and cannot develop into normal offspring⁷. Moreover, in mock experiments, we enucleated 204 oocytes and examined them after fixation and staining⁸: no chromosomes were apparent, suggesting that the efficiency of chromosome removal exceeded 99.99%.

In general, nuclei have previously been transferred either into enucleated, one-cell embryos⁹, or into unfertilized, enucleated oocytes^{10,11} which were then immediately activated, thereby preventing chromosome condensation¹². Mouse embryonic stem¹³ or primordial germ¹⁴ cell nuclei transferred into enucleated oocytes that were then immediately activated produced embryos capable of developing to blastocysts, with a minority of these implanting. In contrast, activation that was delayed for 30 to 60 minutes after the introduction of thymocyte nuclei into enucleated oocytes¹⁵ often resulted in the extrusion of a pseudo-polar body, with a consequently high incidence of hypoauploidy (78%) and with none of the embryos developing beyond the 4-cell stage.

We have shown that a relatively high proportion of enucleated oocytes can develop to morulae/blastocysts and beyond when they are activated after a prolonged delay following injection of adult-

derived, somatic cell nuclei. Indeed, the inclusion of a prolonged interval between nuclear injection and oocyte activation (and suppression of cytokinesis) was apparently beneficial for both pre- and post-implantation development (Tables 1, 3). Although this seems paradoxical after earlier work, prolonged exposure of incoming nuclei to a cytoplasm rich in metaphase promoting factor causes persistent chromosome condensation (in the absence of DNA synthesis) and may facilitate the nuclear changes that are essential for development. We are studying the molecular events attending this latent period with respect to potential epigenetic 'reprogramming' and chromatin repair, *inter alia*. Also, the use of a piezo-impact pipette drive unit^{16,17} may have contributed to a high rate of embryonic development by enabling oocyte and donor nucleus manipulation to be quick and efficient, thereby reducing the trauma to both in comparison with methods using electrofusion, Sendai virus or polyethylene glycol. Furthermore, we minimized the amount of somatic cell cytoplasm introduced into enucleated oocytes, which might otherwise have interfered with the onset of development.

It is unclear why Sertoli and neuronal cell nuclei failed to produce full-term embryos. Although our study does not preclude the possibility that these nuclei (and nuclei from other cell types) may be able to support full-term development, this finding suggests that the G0 status of donor nuclei is not sufficient *per se* to ensure embryonic development. We did not use mural granulosa cells, which differ functionally and in their subsequent fate from cumulus cells¹⁸, and the question is open as to whether their nuclei might prime embryonic development to term. The contrastingly high implantation rate (57–71%) and low fetal (5–16%) and full-term (2–3%) developmental rates (Table 3) indicate that several regulatory morphogenic factors and checkpoints may be involved in the development of post-implantation embryos/fetuses.

Our results suggest that, contrary to previous opinion⁹, mammals can be reproducibly cloned from adult somatic cells. Furthermore, we believe that the success of these experiments in the mouse provides an amenable model with which to evaluate the molecular mechanisms that regulate the reprogramming of somatic cell genomes, genomic imprinting, embryonic genome activation and cell differentiation. □

Methods

The cloning procedure is summarized in Fig. 5.

Isolation of cumulus cells. Female B6D2F1 (C57BL/6 × DBA/2 used in series A and B), B6C3F1 (C57BL/6 × C3H/He used in series C) or B6C3F1 clones produced in series C were induced to superovulate by consecutive injection of eCG and hCG. 13 h after hCG injection, cumulus–oocyte complexes were collected from oviducts and treated in HEPES–CZB medium¹⁹ supplemented with bovine testicular hyaluronidase (0.1% (w/v), 300 U mg⁻¹) to disperse cumulus cells. We selected cumulus cells of modal (>70%) diameter (10–12 µm) for injection. From preliminary experiments, nuclei from cells with smaller or larger diameters (8–9 or 13–15 µm, respectively) seldom supported development of injected eggs beyond the 8-cell stage (data not shown). Following dispersal, cells were transferred to HEPES–CZB containing 10% (w/v) polyvinylpyrrolidone (average *M_r*, 360,000) and kept at room temperature for up to 3 h before injection.

Isolation of Sertoli cells and neurons. Sertoli cells were isolated from the testes of 6-month-old B6D2F1 males as described²⁰, except that HEPES–Ham F-12 medium was used. Manipulation of individual Sertoli cells was done by using a large injection pipette (inner diameter ~10 µm). Neuronal cells were isolated from the cerebral cortex of adult B6D2F1 females. Brain tissue was removed with sterile scissors, quickly washed in erythrocyte-lysing buffer and gently hand-homogenized for several seconds in nucleus isolation medium²¹ at room temperature. Nuclei (7–8 µm in diameter) harbouring a conspicuous nucleolus were individually collected from the resulting suspension using the injection pipette before delivery into a recipient enucleated oocyte.

Enucleation of metaphase II oocytes and donor cell nucleus injection. B6D2F1 oocytes (obtained 13 h after hCG injection of eCG-primed females)

were freed from the cumulus oophorus and held in CZB medium at 37.5 °C under 5% (v/v) CO₂ in air until required. Groups of oocytes (usually 10–15) were transferred into a droplet of HEPES–CZB containing 5 µg ml⁻¹ cytochalasin B, which had previously been placed in the operation chamber on the microscope stage. Oocytes undergoing microsurgery were held with a holding pipette and the zona pellucida 'cored' following the application of several piezo-pulses to an enucleation pipette. The metaphase II chromosome–spindle complex (identifiable as a translucent region) was aspirated into the pipette with a minimal volume of oocyte cytoplasm²². After enucleation of all oocytes in one group (~10 min), they were transferred into cytochalasin B-free CZB and held there for up to 2 h at 37.5 °C, then returned to the microscope stage immediately before further manipulation. Nuclei were removed from their respective somatic cells and gently aspirated in and out of the injection pipette (~7 µm inner diameter) until their nuclei were largely devoid of visible cytoplasmic material. Each nucleus was injected into a separate enucleated oocyte within 5 min of its isolation as described¹⁷.

Oocyte activation. Following somatic-cell nucleus injection, some groups of oocytes were placed immediately in Ca²⁺-free CZB containing both 10 mM Sr²⁺ and 5 µg ml⁻¹ cytochalasin B for 6 h. Additional groups of enucleated oocytes injected with cumulus cell nuclei were left in CZB medium at 37.5 °C under 5% (v/v) CO₂ in air for 1–6 h before activation by Sr²⁺ in the presence of 5 µg ml⁻¹ cytochalasin B. Sr²⁺ treatment activated the oocytes²³, whereas cytochalasin B prevented subsequent polar-body formation and therefore chromosome expulsion. Following activation, all resulting embryos were transferred to Sr²⁺-free, cytochalasin B-free CZB medium and incubation was continued at 37.5 °C under 5% (v/v) CO₂ in air.

Embryo transfer. Where appropriate, 2- to 8-cell embryos or morulae/blastocysts were respectively transferred into oviducts or uteri of foster mothers (CD-1, albino) that had been mated with vasectomized CD-1 males 1 or 3 days previously. Following caesarean section of recipient females at 18.5–19.5 d.p.c., live young were raised by lactating CD-1 foster mothers.

DNA typing. DNA from the following control strains and hybrids was obtained from spleen tissue: C57BL/6J (B6), C3H/HeJ (C3), DBA/2J (D2), B6C3F1 and B6D2F1. DNA from the three cumulus cell donor females (B6C3F1), the three oocyte recipient females (B6D2F1) and the three foster females (CD-1) was prepared from tail-tip biopsies. DNA from the six B6C3F1-derived, cloned offspring was prepared from their associated placentas. For the microsatellite markers *D1Mit46*, *D2Mit102* and *D3Mit49*, primer pairs (MapPairs) were purchased from Research Genetics and typed as described²⁴, except that PCR was carried out for 30 cycles and products were separated by 3% agarose gels (Metaphor) and visualized by ethidium bromide staining. Endogenous ecotropic murine leukaemia provirus DNA sequences (*Env* loci) were identified following hybridization of *Pvu*II-digested genomic DNA to the diagnostic probe, pEc-B4 (ref. 25). Probe labelling, Southern blotting and hybridization procedures have been described²⁶.

Received 23 December 1997; accepted 19 June 1998.

- Campbell, K. H. S., Loi, P., Otaegui, P. J. & Wilmut, I. Cell cycle co-ordination in embryo cloning by nuclear transfer. *Rev. Reprod.* **1**, 40–45 (1996).
- Kono, T. Nuclear transfer and reprogramming. *Rev. Reprod.* **2**, 74–80 (1997).
- Campbell, K. H. S., McWhir, J., Ritchie, W. A. & Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380**, 64–66 (1996).
- Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. S. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810–813 (1997).
- Schuetz, A. W., Whittingham, D. G. & Snowdon, R. Alterations in the cell cycle of mouse cumulus granulosa cells during expansion and mucification *in vivo* and *in vitro*. *Reprod. Fertil. Dev.* **8**, 935–943 (1996).
- Wasserman, P. The biology and chemistry of fertilization. *Science* **235**, 553–560 (1987).
- Epstein, C. J. *The Consequences of Chromosome Imbalance* (Cambridge University Press, 1986).
- Yanagida, K., Yanagimachi, R., Perreault, S. D. & Kleinfeld, R. G. Thermotability of sperm nuclei assessed by microinjection into hamster oocytes. *Biol. Reprod.* **44**, 440–447 (1991).
- McGrath, J. & Solter, D. Inability of mouse blastomere nuclei transferred to enucleated zygotes to support development *in vitro*. *Science* **226**, 1317–1319 (1984).
- Willadsen, S. M. Nuclear transplantation in sheep embryos. *Nature* **320**, 63–65 (1986).
- Collas, P. & Barnes, F. L. Nuclear transplantation by microinjection of inner cell mass and granulosa cell nuclei. *Mol. Reprod. Dev.* **38**, 264–267 (1994).
- Czolowska, R., Modlinski, J. A. & Tarkowski, A. K. Behavior of thymocyte nuclei in non-activated and activated mouse oocytes. *J. Cell Sci.* **69**, 19–34 (1984).
- Tsunoda, T. & Kato, Y. Nuclear transplantation of embryonic stem cells in mice. *J. Reprod. Fertil.* **98**, 537–540 (1993).
- Tsunoda, T., Tokunaga, T., Imai, I. & Uchida, T. Nuclear transplantation of male primordial germ cells in the mouse. *Development* **107**, 407–411 (1989).
- Kono, T., Ogawa, M. & Nakahara, T. Thymocyte transfer to enucleated oocytes in the mouse. *J. Reprod. Dev.* **39**, 301–307 (1993).
- Kimura, Y. & Yanagimachi, R. Intracytoplasmic sperm injection in the mouse. *Biol. Reprod.* **52**, 709–720 (1995).

17. Kimura, Y. & Yanagimachi, R. Mouse oocytes injected with testicular spermatozoa or round spermatids can develop into normal offspring. *Development* **121**, 2397–2405 (1995).
18. Eppig, J., Wigglesworth, K., Pendola, F. & Hira, Y. Murine oocytes suppress expression of luteinizing hormone receptor messenger ribonucleic acid by granulosa cells. *Biol. Reprod.* **56**, 976–984 (1997).
19. Chatot, C. L., Lewis, J. L., Torres, I. & Ziomek, C. A. Development of 1-cell embryos from different strains of mice in CZB medium. *Biol. Reprod.* **42**, 432–440 (1990).
20. Erickson, R. P., Zwigman, T. & Ao, A. Gene expression, X-inactivation, and methylation during spermatogenesis: the case of *Zfx*, *Zfx* and *Zfy* in mice. *Mol. Reprod. Dev.* **35**, 114–120 (1993).
21. Kuretake, S., Kimura, Y., Hoshi, K. & Yanagimachi, R. Fertilization and development of mouse oocytes injected with isolated sperm heads. *Biol. Reprod.* **55**, 789–795 (1996).
22. Kono, T., Sotomaru, Y., Sato, Y. & Nakahara, T. Development of androgenetic mouse embryos produced by *in vitro* fertilization of enucleated oocytes. *Mol. Reprod. Dev.* **34**, 43–46 (1993).
23. Bos-Mikich, A., Whittingham, D. G. & Kones, K. T. Meiotic and Mitotic Ca^{2+} oscillations affect cell composition in resulting blastocysts. *Dev. Biol.* **182**, 172–179 (1997).
24. Dietrich, W. et al. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**, 423–447 (1992).
25. Taylor, B. A. & Rowe, L. A mouse linkage testing stock possessing multiple copies of the endogenous ecotropic murine leukemia virus genome. *Genomics* **5**, 221–232 (1989).
26. Johnson, K. R., Cook, S. A. & Davison, M. T. Chromosomal localization of the murine gene and two related sequences encoding high-mobility-group I and Y proteins. *Genomics* **12**, 503–509 (1992).

Acknowledgements. This study was supported in part by grants from the National Institutes of Health, ProBio America Inc., and fellowships from the Japanese Society for the Promotion of Science (T.W.) and European Molecular Biology Organization (A.C.F.P.). We thank H. Tateno for chromosome analysis, Y. Nakamura for help with DNA fingerprinting, H. Kishikawa, T. Kasai and R. Kleinfeld for assistance in preparing this manuscript, and J. Eppig for help and advice.

Correspondence and requests for materials should be addressed to R.Y.

Defects in somite formation in *lunatic fringe*-deficient mice

Nian Zhang & Thomas Gridley

The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609-1500, USA

Segmentation in vertebrates first arises when the unsegmented paraxial mesoderm subdivides to form paired epithelial spheres called somites^{1,2}. The Notch signalling pathway is important in regulating the formation and anterior–posterior patterning of the vertebrate somite^{3–7}. One component of the Notch signalling pathway in *Drosophila* is the *fringe* gene, which encodes a secreted signalling molecule required for activation of Notch during specification of the wing margin^{8–11}. Here we show that mice homozygous for a targeted mutation of the *lunatic fringe* (*Lfng*) gene, one of the mouse homologues^{12,13} of *fringe*, have defects in somite formation and anterior–posterior patterning of the somites. Somites in the mutant embryos are irregular in size and shape, and their anterior–posterior patterning is disturbed. Marker analysis revealed that in the presomitic mesoderm of the mutant embryos, sharply demarcated domains of expression of several components of the Notch signalling pathway are replaced by even gradients of gene expression. These results indicate that *Lfng* encodes an essential component of the Notch signalling pathway during somitogenesis in mice.

The *Lfng* gene is expressed during somitogenesis in mice in a dynamic pattern that suggests a possible role for *Lfng* in regulating somite formation and establishing somite borders^{12,13}. To analyse the role of the *Lfng* gene during embryogenesis, we constructed a targeting vector that deleted 0.7 kilobases (kb) of genomic sequence encoding the putative signal peptide and proprotein region of the *Lfng* protein, and replaced the deleted sequence with the *lacZ* gene of *Escherichia coli* (Fig. 1a, b). Mice heterozygous for the *Lfng*^{LacZ} mutant allele were viable and fertile. The pattern of RNA expression from the *Lfng*^{LacZ} mutant allele was identical to that of *Lfng* RNA expression, but expression of β -galactosidase protein from the *Lfng*^{LacZ} mutant allele was not useful as a marker for visualizing the normal pattern of *Lfng* expression during somitogenesis owing to the perdurance of the β -galactosidase protein and the dynamic nature of the *Lfng* expression pattern (Fig. 1c–e).

At birth, *Lfng*^{LacZ} homozygous neonates had a shortened trunk with a rudimentary tail (Fig. 2a). Some *Lfng*^{LacZ} homozygous neonates died within a few hours of birth, apparently from respira-

tory difficulties due to malformed rib cages (see below), but other, less severely affected, *Lfng*^{LacZ} homozygotes could survive to adulthood. Analysis of stained skeletal preparations revealed substantial defects in formation of the vertebral column and ribs in the *Lfng*^{LacZ} homozygotes (Fig. 2b, c). The regular metameric pattern of the vertebrae was disrupted along the entire longitudinal axis. The ribs of the *Lfng*^{LacZ} homozygotes were bifurcated and fused, and some ribs were detached from the vertebral column (Fig. 2c).

Lfng^{LacZ} homozygous mutant embryos could be distinguished

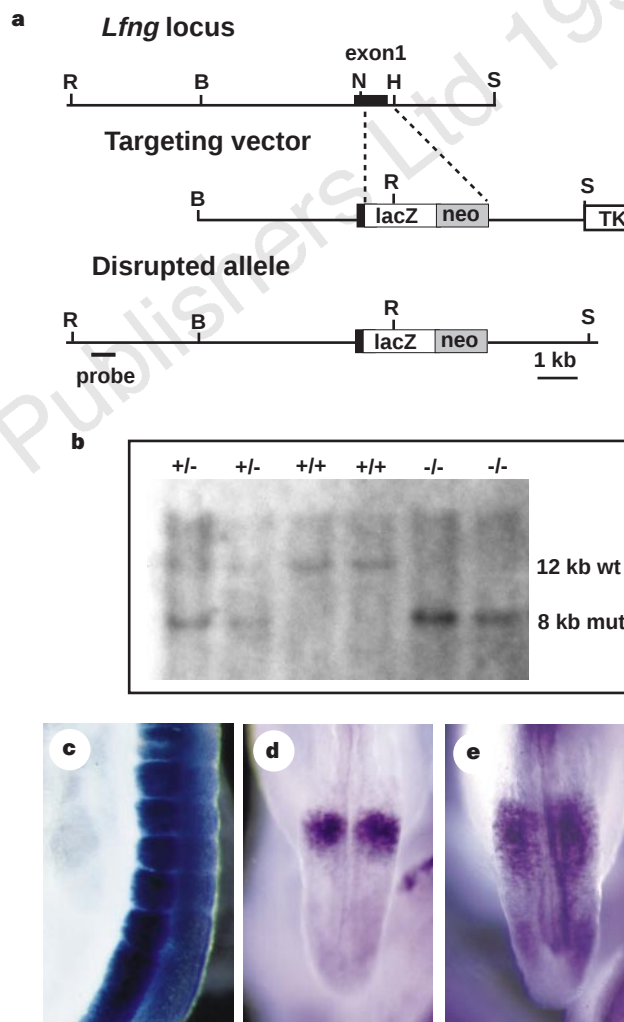


Figure 1 Targeted disruption of the *Lfng* gene. **a**, Targeting scheme. The top line shows the genomic organization of a portion of the *Lfng* gene; the middle line shows the structure of the targeting vector. A 0.7-kb deletion was created which removes most of exon 1, replacing it with the *lacZ* gene and a *neo* cassette. At the bottom is the predicted structure of the *Lfng* locus following homologous recombination of the targeting vector. The probe used for Southern blot analysis in **b** is indicated. Restriction enzymes: B, *Bam*HI; H, *Hind*III; N, *Not*I; R, *Eco*RV; S, *Sal*I. **b**, DNA isolated from embryos of the intercross of *Lfng*^{LacZ}/+ heterozygous mice was digested with *Eco*RV, blotted, and hybridized with the indicated probe. Genotypes of progeny are indicated at the top of each lane. **c**, β -Galactosidase expression in *Lfng*^{LacZ}/+ heterozygous embryos revealed that, unlike *Lfng* RNA, β -galactosidase protein was expressed throughout the rostral presomitic mesoderm and the most recently formed somites. **d**, *In situ* hybridization of a *Lfng*^{LacZ}/+ heterozygous embryo with a *lacZ* probe revealed the same banding pattern in the presomitic mesoderm as is observed with a *Lfng* probe, demonstrating that the constitutive expression observed in **c** is due to perdurance of β -galactosidase. **e**, *In situ* hybridization of a *Lfng*^{LacZ} homozygous mutant embryo with a *lacZ* probe revealed that the stripe of *lacZ* RNA expression is expanded and is more diffuse than in *Lfng*^{LacZ}/+ heterozygous embryos. In **c–e**, posterior is at the bottom. **c**, Sagittal view; **d**, **e**, dorsal views.

from heterozygous and wild-type littermates at any time after embryonic day (E) 8.5, owing to defects in the formation of the somites. The boundaries between somites of the homozygous mutant embryos were less distinct than those of their littermates (Fig. 2d, e), and mutant somites were irregular in size and shape, with some somites being less than half the size of normal (Fig. 2e–j). Remarkably, some somites in the *Lfng^{LacZ}* homozygous mutant embryos were bifurcated (Fig. 2g, h). However, in contrast to other mouse mutants with segmentation defects that are probably caused by perturbations in the Notch signalling pathway, such as mutations in *Dll1* (ref. 3) and *Mesp2* (ref. 5), the *Lfng^{LacZ}* homozygous mutant embryos formed epithelial somites (Fig. 2e–j).

Differentiated derivatives of the somite were also affected in the *Lfng^{LacZ}* mutants. In wild-type embryos at E12.5, a regular metameric pattern of condensing sclerotome cells was seen. In mutant embryos, sclerotome cells condensed, but the regular metameric pattern was not maintained (Fig. 3a, b). Repeating segmental defects were not confined to the derivatives of the somites. Both immunohistochemical and histological analysis revealed fusions of the dorsal root ganglia and disruptions in the patterning and axon trajectories of the spinal nerves (Fig. 3c–f). All of these defects are indicative of the loss of proper anterior–posterior patterning of the somite^{3,5,6,14}.

To examine the effect of the *Lfng^{LacZ}* mutation on the differentiation of somitic cell lineages, we analysed the expression of several molecular markers expressed in somites (Fig. 4). *Myogenin* is

expressed in the myotomes of wild-type embryos in a repeating metameric pattern along the craniocaudal axis¹⁵. In *Lfng^{LacZ}* mutant embryos, *myogenin* was expressed in a repeating metameric pattern, but the stripes of *myogenin* expression were not evenly spaced and frequently fused (Fig. 4b). In wild-type embryos, *Mox1* is expressed at higher levels in the caudal half of the somite¹⁶. In posterior regions of the mutant embryos, the differential expression of *Mox1* between the cranial and caudal compartment was lost, and the gene was expressed at similar levels throughout the somite (Fig. 4d). Similarly, *pax9* expression in the caudal compartment of the sclerotome¹⁷ was lost in posterior regions of the mutant embryos (Fig. 4f). *Uncx4.1* is also expressed in a metameric pattern in the caudal half of the somites of wild-type embryos¹⁸. In mutant embryos, caudal *Uncx4.1* expression was not lost uniformly, but instead showed a disorganized pattern of expression (Fig. 4h). Our results indicate that these different somitic cell lineages differentiate in *Lfng^{LacZ}* mutant embryos, but that anterior–posterior polarity of the somites is disturbed, particularly in posterior regions of the mutant embryos. Similar findings have been reported by Evrard *et al.*¹⁹, although the phenotype may be slightly more severe for their *Lfng* mutant allele. This may be due to differences in the targeting constructs, or to slight differences in the genetic background of the mutant mice.

We then analysed in wild-type and *Lfng^{LacZ}* mutant embryos the expression of several genes encoding ligands and receptors of the Notch signalling pathway (Fig. 5). In wild-type embryos, *Dll3* is expressed in a faint band at the anterior border of the most recently formed somite. The somite in the process of forming exhibits intermediate levels of *Dll3* expression, with clearly demarcated boundaries between the epithelial somite at the anterior side and the region of high *Dll3* expression on the posterior side (ref. 20, and Fig. 5a). We found that the *Notch2* gene is expressed in an almost identical pattern in wild-type embryos, with the exception that the domain of high-level *Notch 2* expression in the presomitic mesoderm did not extend all the way to the caudal end of the embryo (Fig. 5c). In *Lfng^{LacZ}* mutant embryos, both the *Dll3* and the *Notch2*

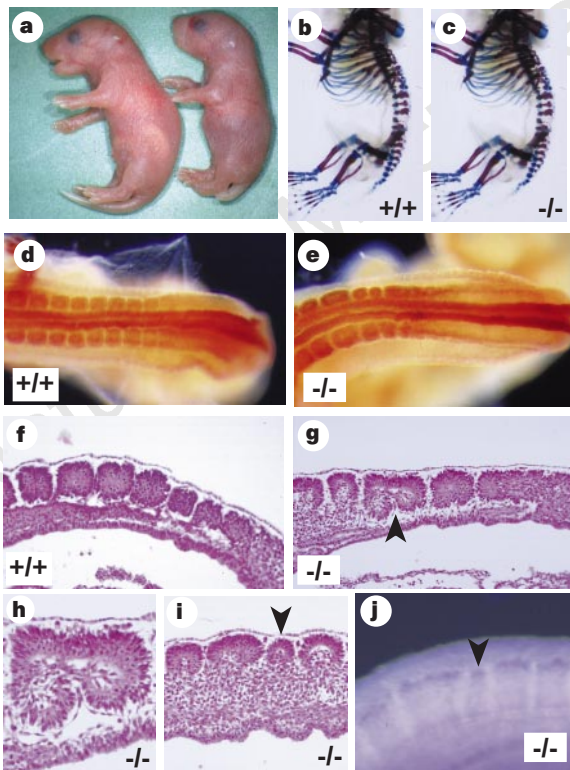


Figure 2 Axial truncation and defects in somite formation in *Lfng^{LacZ}* mutant embryos. **a**, *Lfng^{LacZ}* mutant neonate (on right) has truncation of the anterior–posterior axis both in the trunk and tail. **b–f**, embryo genotypes are indicated. **b**, **c**, Skeletal preparations at E18 stained for bone and cartilage. **d**, **e**, Whole-mount immunohistochemistry with an anti-NCAM antibody at E9.5. Somites of the *Lfng^{LacZ}* mutant embryo are irregular in shape and size, and have less distinct borders than those of the wild-type embryo. **f**, **g**, Parasagittal sections at E9.5. Somites in the *Lfng^{LacZ}* mutant embryo are irregular in size. Arrowhead, a bifurcated somite. **h–j**, Examples of abnormal somites in *Lfng^{LacZ}* mutant embryos at E9.5. **h**, Higher-power view of the indicated (arrowhead) somite in **g**. **i**, **j**, Two very small somites (arrowheads) adjacent to abnormally large somites.

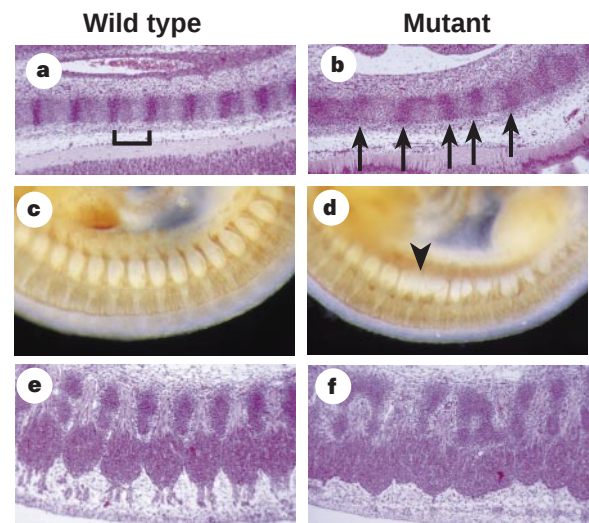


Figure 3 Sclerotome and peripheral nervous system defects in *Lfng^{LacZ}* mutant embryos. **a**, **b**, Parasagittal sections of embryos at E12.5. In the wild-type embryo (**a**), the sclerotome exhibits regularly alternating regions of high and low cell density (bracket) which will form the vertebral bodies and intervertebral discs. In the *Lfng^{LacZ}* mutant embryo (**b**), the metameric pattern of high and low cell density is lost (sclerotome condensations are indicated by arrows in **b**). **c**, **d**, Whole-mount immunohistochemistry with an anti-neurofilament monoclonal antibody at E10.5. The *Lfng^{LacZ}* mutant embryo has fusions of the spinal ganglia and disruptions of the spinal nerves (arrowhead in **d**). **e**, **f**, Parasagittal sections of embryos at E12.5 show fused dorsal root ganglia in the *Lfng^{LacZ}* mutant (**f**).

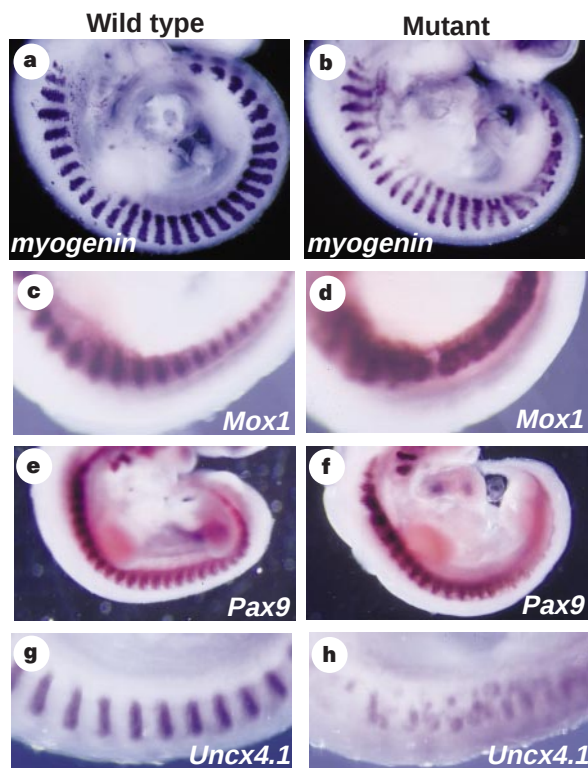


Figure 4 Expression of somite lineage markers in wild-type and *Lfng*^{LacZ} mutant embryos. Whole-mount *in situ* hybridization with the indicated probes to wild-type and mutant embryos isolated at E9.5. In mutant embryos, the stripes of *myogenin* expression frequently fuse (**b**). For both *Mox1* and *Pax9*, the normal anterior-posterior polarity of gene expression is lost in posterior regions of the mutant embryos (**d**, **f**). Anterior is towards the left in all panels.

expression patterns no longer showed sharply demarcated boundaries between the domains of different levels of expression, but instead exhibited a smooth gradation in expression from the presomitic region into the region of formed somites (Fig. 5b, d). *Dll1* is expressed in both the presomitic mesoderm and in the posterior region of the formed somite²¹. In *Lfng*^{LacZ} mutant embryos, the posterior band of *Dll1* expression in the formed somite was lost in most somites, although occasional somites retained the posterior *Dll1* expression stripe (Fig. 5f). *Dll1* expression in the rostral presomitic mesoderm of the mutant embryos was also more diffuse than in controls (Fig. 5e, f). Likewise, the sharp boundary of the *Notch1* expression domain with the most recently formed somite was diffuse in *Lfng*^{LacZ} mutant embryos (Fig. 5g, h). In contrast, the stripe of expression of the *Serrate* homologue *Jag1* at the border of the forming somite^{12,22} was unaffected in *Lfng*^{LacZ} mutant embryos (Fig. 5i, j).

We propose that the phenotypic effects of the *Lfng*^{LacZ} mutation on somitogenesis are caused by perturbations of the Notch signalling pathway. Current models for the function of the fringe protein during formation of the wing margin in *Drosophila* indicate that, at the post-transcriptional level, fringe protein positively regulates signalling by the Delta protein and negatively regulates signalling by Serrate in the wing imaginal disc^{9–11}. This regulation of ligand signalling forms part of a mechanism, which also includes a positive-feedback loop between Delta- and Serrate-expressing cells^{23,24}, that acts to restrict Notch activation to cells along the dorsal-ventral compartment boundary^{9–11}. We find that loss of *Lfng* function leads to loss in the rostral presomitic mesoderm of the sharply demarcated domains of expression of several genes encoding members of the Notch signalling pathway. Feedback mechanisms similar to those occurring at the dorsal-ventral compartment

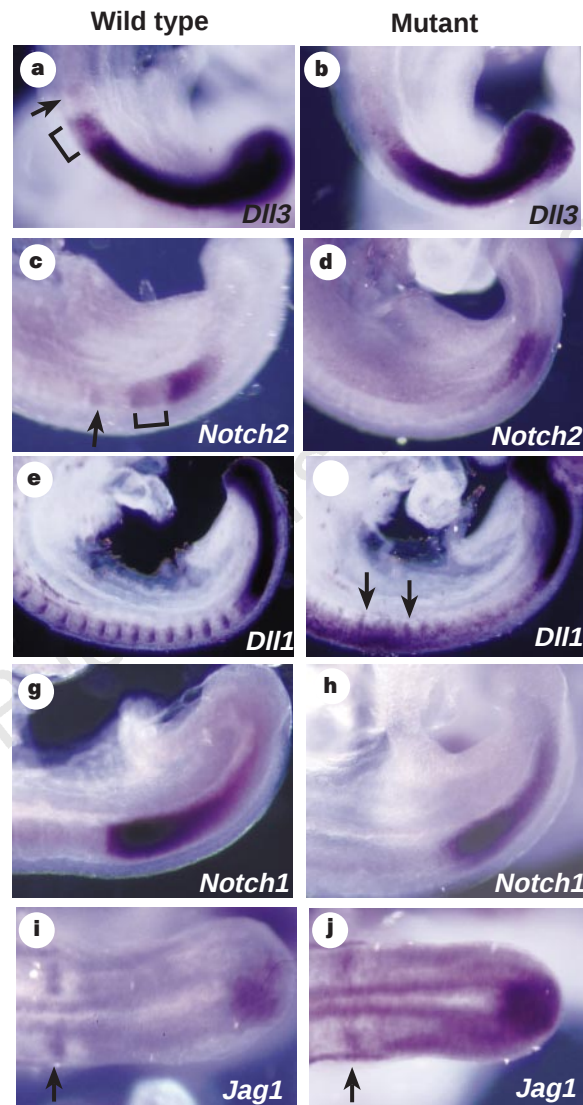


Figure 5 Expression of markers of the Notch signalling pathway in wild-type and *Lfng*^{LacZ} mutant embryos. Whole-mount *in situ* hybridization was done with the indicated probes to wild-type and mutant embryos isolated at E9.5. In wild-type embryos, both *Dll3* (**a**) and *Notch2* (**c**) are expressed in clearly demarcated domains at the anterior side of the most recently formed somite (arrow), in the forming somite (bracket), and in the presomitic mesoderm. In mutant embryos, both *Dll3* (**b**), and *Notch2* (**d**) are expressed in an even gradient, without clearly demarcated domains of expression. *Dll1* expression in the posterior region of the formed somites of mutant embryos is retained in only a few somites (arrows in **f**), and *Dll1* expression in the rostral presomitic mesoderm of the mutants is more diffuse than in wild-type embryos. In mutant embryos, the boundary of the *Notch1* expression domain with the most recently formed somite is more diffuse than in wild-type embryos (**g**, **h**). The *Serrate* homologue *Jag1* is expressed at low levels in the posterior region of the forming somite (arrow) in both wild-type (**i**) and mutant (**j**) embryos. Anterior is towards the left in all panels.

boundary of the wing disc may be involved in generating these sharply demarcated expression domains in the rostral presomitic mesoderm. We propose that the *Lfng* protein and one or both of the Notch ligands *Dll1* and *Dll3* function together in the presomitic mesoderm to mediate the cell-cell interactions that allocate populations of cells to the nascent somites, and that the primary defect in the *Lfng*^{LacZ} mutants is a failure to transduce the signal mediated by the Notch ligands *Dll1* and/or *Dll3*. We further propose that failure of signalling mediated by these Delta homologues leads to the loss of the sharply demarcated domains of gene expression, which may

alter the positioning of the somite boundaries and disturb the allocation of cells to the nascent somite. Thus, as epithelial somites form in the *Lfng*^{LacZ} mutant embryos, varying numbers of cells are allocated to the nascent somite, leading to the observed irregularities in size and shape.

We have shown that, in addition to defects in the initial formation of somites from the rostral presomitic mesoderm, anterior-posterior patterning of the formed somites is disturbed in *Lfng*^{LacZ} mutants. Disruption of anterior-posterior somite patterning may be a secondary consequence of the presomitic mesoderm defects, or may reflect another role of *Lfng* in directly establishing somite polarity in the formed somites. Further work, including the removal of *Lfng* function by gene targeting either only in the presomitic region or only in the posterior part of the formed somite, will be necessary to distinguish between these alternatives. □

Methods

Gene targeting. Mouse *Lfng* genomic clones were isolated from a genomic λ phage 129/Sv library (Stratagene). To construct the targeting vector, a 4-kb *HindIII*-*NotI* fragment was subcloned upstream and a 2.5-kb *HindIII*-*Sall* fragment was subcloned downstream of cassette containing the *lacZ* gene and a *neo*-selection cassette. This resulted in the deletion of a 0.7-kb genomic fragment containing most of exon 1; we refer to this mutant allele as *Lfng*^{LacZ}. Electroporation of embryonic stem (ES) cells, selection and blastocyst injections have been described²⁵. Germline transmission of the *Lfng*^{LacZ} mutant allele was obtained for four targeted clones.

Histology, skeletal analysis and *in situ* hybridization. Embryos were dissected from timed matings of *Lfng*^{LacZ}/+ heterozygotes, and yolk-sac DNA was prepared and genotyped by allele-specific PCR. Embryos for histological analysis were fixed in Bouin's fixative and were stained with haematoxylin and eosin. Whole-mount *in situ* hybridization and staining of skeletal preparations with alcian blue/alizarin red were done as described^{26,27}, as was whole-mount immunohistochemistry²⁵.

Received 4 March; accepted 17 June 1998.

- Keynes, R. J. & Stern, C. D. Mechanisms of vertebrate segmentation. *Development* **103**, 413–429 (1988).
- Gossler, A. & Hrabé de Angelis, M. Somitogenesis. *Curr. Top. Dev. Biol.* **38**, 225–287 (1997).
- Hrabé de Angelis, M., McIntyre II, J. & Gossler, A. Maintenance of somite borders in mice requires the *Delta* homologue *Dll1*. *Nature* **386**, 717–721 (1997).
- Jen, W.-C., Wettstein, D., Turner, D., Chitnis, A. & Kintner, C. The Notch ligand, X-Delta-2, mediates segmentation of the paraxial mesoderm in *Xenopus* embryos. *Development* **124**, 1169–1178 (1997).
- Saga, Y., Hata, N., Koseki, H. & Taketo, M. Mesp2: a novel mouse gene expressed in the presegmented mesoderm and essential for segmentation initiation. *Genes Dev.* **11**, 1827–1839 (1997).
- Wong, C. et al. Presenilin 1 is required for *Notch1* and *Dll1* expression in the paraxial mesoderm. *Nature* **387**, 288–292 (1997).
- Shen, J. et al. Skeletal and CNS defects in Presenilin-1-deficient mice. *Cell* **89**, 629–639 (1997).
- Irvine, K. D. & Wieschaus, E. *fringe*, a boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during *Drosophila* wing development. *Cell* **79**, 595–606 (1994).
- Fleming, R. J., Gu, Y. & Hukriede, N. A. Serrate-mediated activation of Notch is specifically blocked by the product of the gene *fringe* in the dorsal compartment of the *Drosophila* wing imaginal disc. *Development* **124**, 2973–2981 (1997).
- Panin, V. M., Papayannopoulos, V., Wilson, R. & Irvine, K. D. *Fringe* modulates Notch-ligand interactions. *Nature* **387**, 908–912 (1997).
- Blair, S. S. Limb development: Marginal Fringe benefits. *Curr. Biol.* **7**, R686–R690 (1997).
- Cohen, B. et al. *Fringe* boundaries coincide with Notch-dependent patterning centres in mammals and alter Notch-dependent development in *Drosophila*. *Nature Genet.* **16**, 283–288 (1997).
- Johnston, S. H. et al. A family of mammalian *fringe* genes implicated in boundary determination and the Notch pathway. *Development* **124**, 2245–2254 (1997).
- Tannahill, D., Cook, G. M. W. & Keynes, R. J. Axon guidance and somites. *Cell Tissue Res.* **290**, 275–283 (1997).
- Montarras, D. et al. Developmental patterns in the expression of Myf5, MyoD, myogenin, and MRF4 during myogenesis. *New Biol.* **3**, 592–600 (1991).
- Candia, A. F. et al. Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development* **116**, 1123–1136 (1992).
- Neubüser, A., Koseki, H. & Balling, R. Characterization and developmental expression of Pax9, a paired-box-containing gene related to Pax1. *Dev. Biol.* **170**, 701–716 (1995).
- Neidhardt, L. M., Kispert, A. & Herrmann, B. G. A mouse gene of the paired-related homeobox class expressed in the caudal somite compartment and in the developing vertebral column, kidney and nervous system. *Dev. Genes Evol.* **207**, 330–339 (1997).
- Evraud, Y. A., Lun, Y., Aulehla, A., Gan, L. & Johnson, R. L. *lunatic fringe* is an essential mediator of somite segmentation and patterning. *Nature* **394**, 377–381 (1998).
- Dunwoodie, S. L., Henrique, D., Harrison, S. M. & Beddington, R. S. Mouse *Dll3*: a novel divergent *Delta* gene which may complement the function of other *Delta* homologues during early pattern formation in the mouse embryo. *Development* **124**, 3065–3076 (1997).
- Bettenhausen, B., Hrabé de Angelis, M., Simon, D., Guenet, J.-L. & Gossler, A. Transient and restricted expression during mouse embryogenesis of *Dll1*, a murine gene closely related to *Drosophila Delta*. *Development* **121**, 2407–2418 (1995).
- Mitsiadis, T. A., Henrique, D., Thesleff, I. & Lendahl, U. Mouse *Serrate-1* (*Jagged1*): Expression in the

- developing tooth is regulated by epithelial-mesenchymal interactions and fibroblast growth factor-4. *Development* **124**, 1473–1483 (1997).
- de Celis, J. F., Garcia-Bellido, A. & Bray, S. J. Activation and function of *Notch* at the dorsal-ventral boundary of the wing imaginal disc. *Development* **122**, 359–369 (1996).
 - Doherty, D., Feger, G., Younger-Shepherd, S., Jan, L. Y. & Jan, Y. N. Delta is a ventral to dorsal signal complementary to Serrate, another Notch ligand, in *Drosophila* wing formation. *Genes Dev.* **10**, 421–434 (1996).
 - Swiatek, P. & Gridley, T. Perinatal lethality and defects in hindbrain development in mice homozygous for a targeted mutation of the zinc finger gene *Krox20*. *Genes Dev.* 2071–2084 (1993).
 - Wilkinson, D. G. In *In Situ Hybridization: A Practical Approach* (ed. Wilkinson, D. G.) 75–83 (IRL, Oxford, 1992).
 - Martin, J. F., Bradley, A. & Olson, E. N. The paired-like homeobox gene *MHox* is required for early events of skeletogenesis in multiple lineages. *Genes Dev.* **9**, 1237–1249 (1995).

Acknowledgements. We thank R. Johnson for exchanging unpublished data; C. Norton for doing the blastocyst injections; A. Gossler, R. Jiang and T. O'Brien for comments on the manuscript; and R. Balling, R. Beddington, A. Gossler, B. Herrmann, E. Olson and C. Wright for probes. This work was supported by grants from the NIH and the March of Dimes Foundation to T.G., and by a grant from the National Cancer Institute to The Jackson Laboratory.

Correspondence and requests for materials should be addressed to T.G. (e-mail: gridley@jax.org).

lunatic fringe is an essential mediator of somite segmentation and patterning

Yvonne A. Evraud*, Yi Lun*, Alexander Aulehla, Lin Gan & Randy L. Johnson

Department of Biochemistry and Molecular Biology and Program in Genes and Development, Box 117, University of Texas, M. D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, Texas 77030, USA

* These authors contributed equally to this work.

The gene *lunatic fringe* encodes a secreted factor with significant sequence similarity to the *Drosophila* gene *fringe*^{1–5}. *fringe* has been proposed to function as a boundary-specific signalling molecule in the wing imaginal disc, where it is required to localize signalling activity by the protein Notch to the presumptive wing margin^{3,6}. By targeted disruption in mouse embryos, we show here that *lunatic fringe* is likewise required for boundary formation. *lunatic fringe* mutants fail to form boundaries between individual

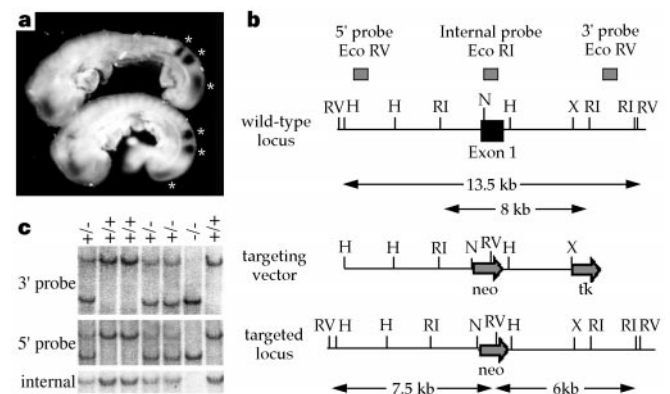


Figure 1 Expression and targeting of *lunatic fringe*. **a**, Whole-mount *in situ* hybridization of 8.75-d.p.c. embryos using a *lunatic fringe* antisense probe. In the somitic mesoderm, several bands of *lunatic fringe* transcripts are present in the segmental plate before somite formation (asterisks). Significant variation is seen in the location of *lunatic fringe* transcripts in the segmental plate. **b**, Targeting strategy. A replacement vector designed to delete the first coding exon of *lunatic fringe* was constructed. The approximate location of probes used (see bars above the wild-type restriction map) and the sizes of restriction fragments used to assess correct targeting are shown. **c**, Southern analysis of DNA extracted from tails of pups derived from heterozygous intercrosses. For each probe (5', 3' and exon 1), the expected restriction fragments were visualized for each genotype. Abbreviations: RV, *EcoRV*; H, *HindIII*; RI, *EcoRI*; N, *NotI*; X, *XhoI*.

somites, the initial segmental unit of the vertebrate trunk. In addition, the normal alternating rostral–caudal pattern of the somitic mesoderm is disrupted, suggesting that intersomitic boundary formation and rostral–caudal patterning of somites are mechanistically linked by a process that requires *lunatic fringe* activity. As a result, the derivatives of the somitic mesoderm, especially the axial skeleton, are severely disorganized in *lunatic fringe* mutants. Taken together, our results demonstrate an essential function for a vertebrate *fringe* homologue and suggest a model in which *lunatic fringe* modulates Notch signalling in the segmental plate to regulate somitogenesis and rostral–caudal patterning of somites simultaneously.

Of the three murine *fringe* homologues (*manic*, *radical* and *lunatic fringe*), only *lunatic fringe* is expressed in the presomitic mesoderm^{1,2}. Within the segmental plate (Fig. 1a), where somites form from a coordinated mesenchymal-to-epithelial transition^{7,8}, three distinct bands of *lunatic fringe* messenger RNA are present at 8.75 days post coitum (d.p.c.). Apart from the somitic mesoderm, *lunatic fringe* is expressed in several other tissues during development, including rhombomeres 3 and 5, the developing ear, the retina and the spinal cord^{1,2}.

To determine the essential requirements for *lunatic fringe* in murine development, we generated a targeted mutation that removes the entire first coding exon (Fig. 1b). This deletion eliminates the translational start site, the signal sequence and a putative processing site, and therefore is predicted to inactivate the *lunatic fringe* protein product. Correct targeting was confirmed by Southern-blot analysis (Fig. 1c). *lunatic fringe* mRNA is not detected in homozygous mutants, suggesting that this allele is a null mutation. *lunatic fringe* heterozygotes appear normal and were fertile. In contrast, homozygous mutant animals had a reduced viability at birth and before weaning; surviving animals were of an obvious phenotype, with truncated tails and shortened body axes (Fig. 2a).

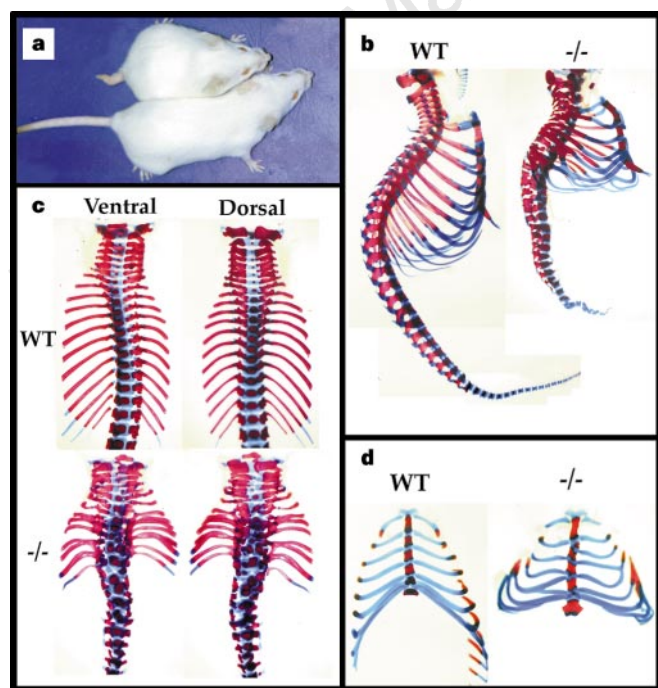


Figure 2 Skeletal phenotype of *lunatic fringe* mutant mice. **a**, Photograph of 8-week-old wild-type and mutant littermates. Mutant animals have a very small and bent tail and their body axis is shortened from the neck to the rump. **b–e**, Alcian blue/Alizarin red skeletal preparations of newborn wild-type and mutant mice. **b**, Side view with limbs removed (WT, wild type; $-/-$, homozygous mutant). **c**, Dorsal and ventral views of the cervical, thoracic and lumbar regions, with the ventral ribs and sternum removed. **d**, Dissected ventral ribs and sternum.

The external appearance of *lunatic fringe* mutants suggested that skeletal abnormalities were present. To evaluate the nature and severity of these defects, skeletal preparations were made of mutant ($n = 10$) and wild-type ($n = 10$) newborn littermates (Fig. 2b–d). The axial skeleton of *lunatic fringe* mutants is severely perturbed, with numerous vertebral and rib fusions and incompletely formed vertebrae. The skeletal anomalies seen in newborn animals are preceded by segmentation defects in the paraxial mesoderm which gives rise to the axial skeleton and in tissues that invade the paraxial mesoderm (Fig. 3a–e). Both the normal alternating density of the sclerotome and the regular spacing of the dorsal root ganglia are perturbed in *lunatic fringe* mutants.

Parasagittal sections of 9.5 d.p.c. embryos demonstrate that mice homozygous for *lunatic fringe* fail to form the initial segmental unit of the paraxial mesoderm, the epithelial somite (Fig. 3b, d). (Other mice homozygous for the *lunatic fringe* allele do not completely lack epithelial somite⁹: the reason for this difference may be due to

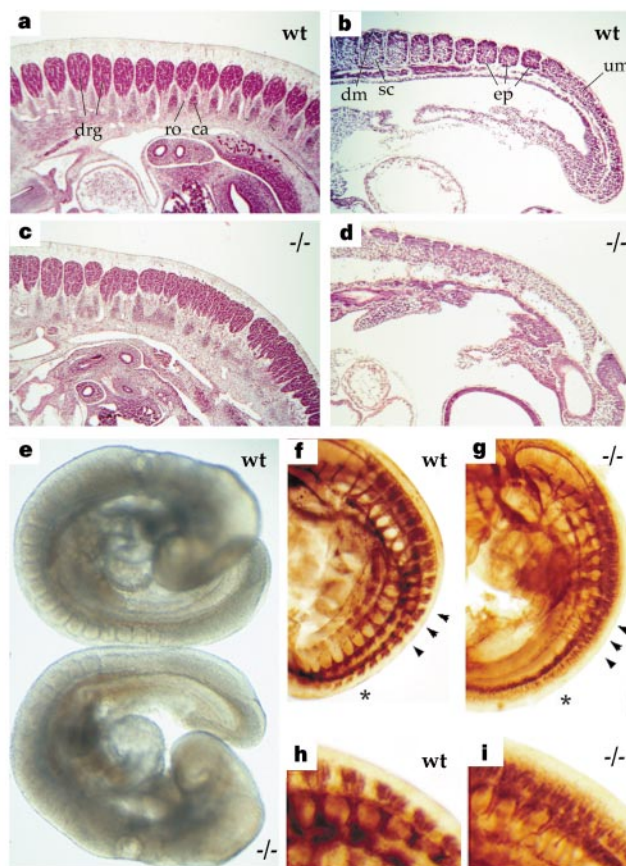


Figure 3 Gross, histological, and immunohistochemical analysis of segmentation in wild-type and *lunatic fringe* mutant embryos. **a, c**, Parasagittal sections of the trunk region of 11.5-d.p.c. wild-type and mutant embryos. Note the regular spacing of dorsal root ganglia (drg) in wild-type embryos and the alternating density of the rostral (ro) and caudal (ca) sclerotome. In the mutants, the drg spacing is irregular and often fused. The rostral–caudal density of the sclerotome is also irregular. **b, d**, Parasagittal sections of the trunk/tail region of 9.0-d.p.c. wild-type and mutant embryos. In wild-type embryos, the epithelial somites (ep) form from the unsegmented presomitic mesoderm (um) and differentiate into sclerotome (sc) and dermomyotome (dm). In *lunatic fringe* mutants, epithelial somites do not form and the dermomyotome size and spacing are irregular. **e**, Photomicrograph of freshly dissected 9.0-d.p.c. embryos. The segmentation defect can be seen in the tail region of the mutant (lower) embryo. **f–i**, Anti-neurofilament staining with the 2H3 monoclonal antibody in 10.0-d.p.c. embryos. The periodic staining pattern of wild-type embryos (arrowheads in **f, g**) is disrupted in the mutants. These regions are shown at higher magnification in **h** and **i**. The onset of 2H3 expression also appears to be delayed (asterisks in **f, g**).

specific differences in genetic background or in the design of the targeting vector.) Despite the apparent absence of epithelial somites in our homozygous mutants, differentiation of the somitic mesoderm occurs on schedule, with formation of an epithelial dermomyotome and sclerotome (Fig. 3b, d, and see below). In addition, the dermomyotome appears to be divided into segmental units, but the regular spacing of these units seems to be disrupted (Fig. 3b, d, e). However, the ventral somitic mesoderm is apparently devoid of segmental information, as assayed by molecular markers (see below).

To confirm and extend our histological analysis, we compared the expression of several presomitic and somitic markers, including *pax1*, *pax3* and *myogenin*^{10–12} in wild-type and *lunatic fringe* mutant embryos (Fig. 4, and data not shown). Expression of *pax1* in *lunatic fringe* is confined to the sclerotome, as in wild-type embryos, but its expression is no longer segmented (Fig. 4c, m, o). Similar, but less pronounced, effects are observed for *pax-3* and *myogenin*, the expression of which is confined to the dermomyotome and myotome, respectively (Fig. 4a, b).

lunatic fringe mutant embryos have defects in the rostral–caudal polarity of their somitic mesoderm, as suggested by histological analysis at 11.5 d.p.c. (Fig. 3a, c). To evaluate the initial rostral–caudal somite pattern of *lunatic fringe* mutant embryos, we examined the expression of *uncx4.1*, which is normally confined to the caudal half-somite¹³ (Fig. 4d, n). In *lunatic fringe* mutants, *uncx4.1* message is initially found in a patchy pattern, intermingled with cells that do not express *uncx4.1* (Fig. 4d, p). Similar results were obtained for *delta-1* (ref. 14; Fig. 4e, q, s). A defect in patterning of the rostral half-somite was revealed by staining with the anti-neurofilament antibody 2H3 (ref. 15; Fig. 3f–i). In contrast to the wild-type periodic 2H3 staining pattern in the rostral portion of each somite, *lunatic fringe* mutants exhibit a fusion of 2H3 immunoreactivity along the posterior flank. Taken together, these observations suggest that the somitic mesoderm of *lunatic fringe* mutants is initially neither exclusively rostral nor caudal, but a mixture of the two, and that some of the defects present in *lunatic fringe* mutants relate to an inability to define caudal and rostral half-somites spatially.

Segmentation and rostral–caudal patterning appears to be less severely affected in anterior somites of *lunatic fringe* mutant embryos (Figs 3c, 4d). Two mechanisms could account for this. First, segmentation could be differentially affected in anterior versus posterior somites. This is unlikely to be the case from 9.0–11.5 d.p.c. because we observe an equivalent inability to form epithelial somites at 9.0, 10.5 and 11.5 d.p.c. (Fig. 3, and data not shown). A second possibility is that initial defects in rostral–caudal pattern can be corrected by a mechanism involving cell sorting in the somitic mesoderm. Once formed, the caudal half-somites and the rostral half-somites are differentially adhesive and this property is thought to maintain the segmental organization of the paraxial mesoderm¹⁶. As the paraxial mesoderm of *lunatic fringe* mutants appears to be a mixture of rostral and caudal cells, we propose that they can sort out into rostral and caudal clusters. However, as the segmental organization of the paraxial mesoderm is disrupted, these clusters are irregularly patterned.

lunatic fringe is a homologue of the *Drosophila* gene *fringe*, which modulates Notch signalling at the presumptive wing margin⁶. To see whether *lunatic fringe* might also regulate Notch activity in the presomitic mesoderm of murine embryos, we examined the expression of a Notch target (*hes-5*; ref. 17, 18) in *lunatic fringe* mutants. In the segmental plate of 9.0–9.5 d.p.c. embryos ($n = 36$), *hes-5* transcripts are localized in two bands at the rostral end of the segmental plate, with the rostral band corresponding approximately to the future somite boundary (Fig. 4f, g, r). In contrast, *hes-5* expression is reduced ($n = 9$) or lost completely ($n = 1$) in the paraxial mesoderm of *lunatic fringe* mutants (Fig. 4h, i, t). Because *hes-5* expression is also downregulated in the paraxial mesoderm of *notch-1* and *RBP-Jk* mutants¹⁸, our results indicate that *lunatic fringe* is required for proper Notch-mediated expression of *hes-5* in the segmental plate. The similarity of the *lunatic fringe* and Notch-pathway^{19–21} mutant phenotypes also supports the idea that *lunatic fringe* modulates Notch signalling in the presomitic mesoderm.

Unlike *hes-5*, the presomitic expression of *notch* (*notch-1*, 2; ref. 22) and of its ligands (*delta-1*, *delta-3*, *serrate-1*; refs 14, 23, 24) are not grossly affected in *lunatic fringe* mutants, although their rostral

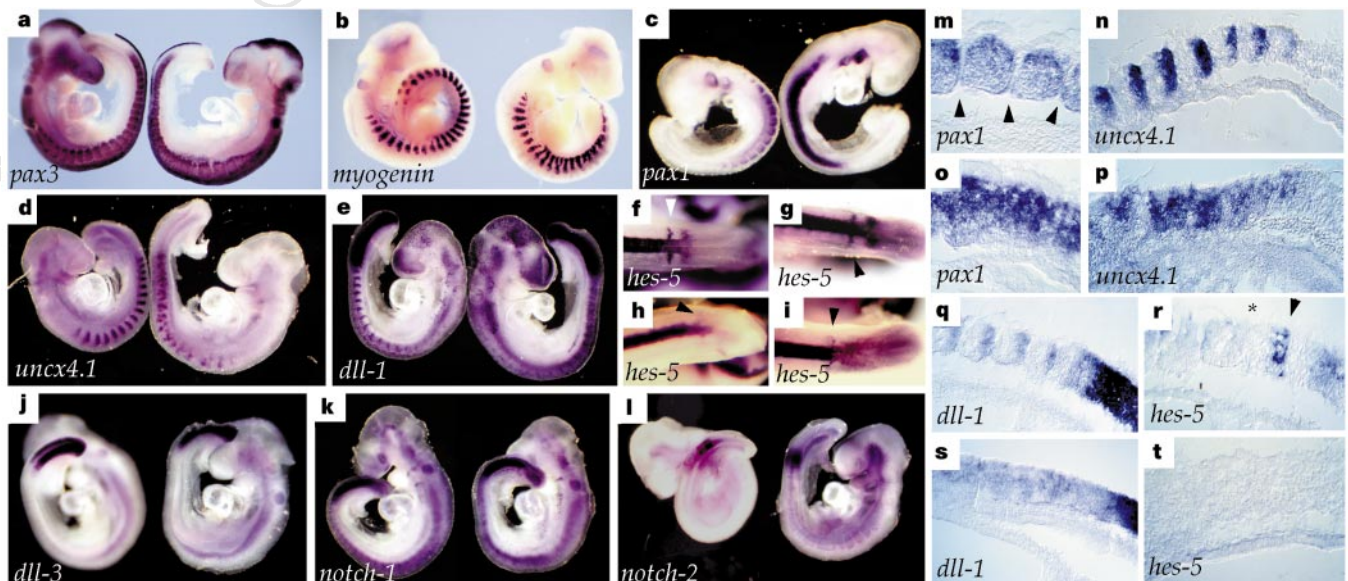


Figure 4 Gene expression is altered in the somitic mesoderm of *lunatic fringe* mutants at 9–9.5 d.p.c. **a–l**, Whole-mount *in situ* hybridization with antisense riboprobes. The probes used are indicated at the lower left. In **a–e** and **j–l**, the wild-type embryo is on the left and the homozygous mutant on the right. For *hes-5*, **f, g** show results from wild-type, and **h, i** from mutants. Note the narrow rostral band of expression in the presomitic mesoderm of wild-type embryos (**f, g**) and its absence or down regulation in mutants (**h, i**). **m–t**, Cryosections of whole-mount

in situ hybridizations **m, n, q, r**, wild type, **o, p, s, t**, mutants. Note the intersomitic boundaries between *pax1* expression domains in wild-type embryos (arrowheads in **m**) and their absence in the mutants (**o**). Also note the location of the rostral *hes-5* band of expression (arrowhead in **r**) corresponding to the presumptive posterior somite and/or the presumptive intersomitic boundary. This band is missing or downregulated in mutants.

and caudal boundaries are more diffuse (Fig. 4, and data not shown). Moreover, the expression of *delta-1* and *hes-5* in the central nervous system is unaffected in *lunatic fringe* mutants, suggesting that Notch signalling in neuroblasts is not significantly altered by loss of *lunatic fringe*. Overlapping expression of *manic* and/or *radical fringe*^{1,2} may compensate for the lack of *lunatic fringe* in these and may be other tissues.

Our results show that *lunatic fringe* is required for both somite segmentation and rostral-caudal patterning. In principle, several distinct mechanisms that differ in where *lunatic fringe* activity is required, and whether segmentation or rostral-caudal patterning is the primary defect, could account for these observations (Fig. 5). One possibility is that *lunatic fringe* may initiate or refine a metamereric prepattern in the segmental plate (Fig. 5a), as has been proposed for Notch signalling²⁵. Disruption of this metamereric prepattern would simultaneously interfere with the organization of cells into segmental units and with rostral-caudal patterning within each segmental unit. Alternatively, an initial defect in rostral-caudal patterning could cause secondary defects in segmentation.

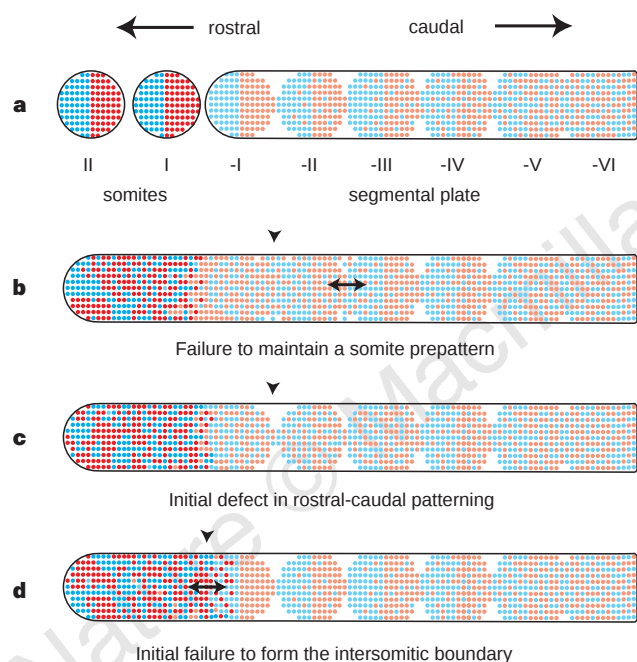


Figure 5 Models for *lunatic fringe* function during somite segmentation and rostral-caudal patterning. **a**, Normal segmental plate and somite development. Somites form at the rostral end of the segmental plate and are divided into rostral (dark blue) and caudal (red) cells. Cells in the segmental plate are progressively organized in a rostral (light blue)-caudal (orange) somitomer prepattern. As individual segmental plate cells are committed to neither a single somite nor to rostral or caudal fates until just before somite formation, any prepattern cannot define precise somite or half-somite units until immediately before overt segmentation. **b**, Effect of a failure to maintain the somitomer prepattern: cells from adjacent somitomes mix (double arrow) and disrupt the rostral-caudal prepattern. Intersomitic boundaries do not form because the somitomer prepattern is not maintained. The net result is an absence of epithelial somites and a mixing of rostral and caudal cells within the extended unsegmented mesoderm. **c**, Effect of an initial defect in rostral-caudal patterning: the somitomer prepattern of the segmental plate is unaffected but cells of the segmental plate have a disorganized rostral-caudal prepattern. If an interface between presumptive rostral and caudal cells were necessary at the presumptive intersomitic boundary (arrowhead), somite boundaries would not form. **d**, Effect of a failure to form the intersomitic boundary: the somitomer and rostral-caudal prepatterns are unaffected. At the position of normal intersomitic boundary formation (arrowhead), caudal cells and presumptive rostral cells mix, leading to secondary defects in rostral-caudal patterning.

Intersomitic boundaries form at the borders of alternating rostral and caudal half-somites during normal development and can also be created experimentally by grafting of half-somites¹⁶. Therefore, an inability properly to pattern rostral and/or caudal half-somites could result in a failure to form the intersomitic boundary. Finally, a defect in intersomitic boundary formation may be the primary cause of the *lunatic fringe* mutant phenotype. In *Drosophila*, juxtaposition of *fringe*-expressing and non-expressing cells at the dorsal-ventral compartment interface specifies the location of the wing margin in a process that requires Notch signalling^{3,6}. The presence of *lunatic fringe*-expressing and non-expressing cells at the presumptive intersomitic boundary^{1,2} and the failure to form the intersomitic boundary in *lunatic fringe* mutants indicates that a similar mechanism may be operating in the presomitic mesoderm of vertebrates. These three models are not necessarily mutually exclusive, and *lunatic fringe* may be required for multiple patterning processes in the presomitic mesoderm. Irrespective of precisely how *lunatic fringe* works, our results highlight its importance in modulating somite segmentation and patterning and provide evidence for an essential function for *fringe* homologues in vertebrate boundary formation. □

Methods

Isolation and expression analysis of murine *lunatic fringe*. Using human expressed-sequence tag (EST) cDNAs with sequence similarity to *Drosophila fringe* (from E. Laufer), an 8.5-d.p.c. whole-embryo cDNA library (from B. Hogan) was screened under conditions of reduced stringency. Several strongly hybridizing clones were isolated, and one (pMFR1, 2.2 kb) was selected for further study. Sequence analysis indicated that it contained the entire open reading frame of murine *lunatic fringe*^{1,2}. The entire insert from pMFR1 was used for whole-mount *in situ* hybridization according to ref. 26.

Targeted disruption of *lunatic fringe*. A genomic library (from P. Hasty) was screened with the insert from pMFR1 and several clones were identified, one of which contained the first coding exon of *lunatic fringe*. A targeting vector (pMFRTV1) was constructed to delete the first coding exon. Electroporation of this construct into AB1 embryonic stem (ES) cells²⁷ and analysis of the doubly resistant colonies by Southern hybridization gave correctly targeted clones with ~30% targeting frequency. Chimaeras were generated by injecting ES cells into C57Bl/6 blastocysts and two independent lines passed the targeted allele through the germ line.

Analysis of *lunatic fringe* mutants. All analysis was conducted on a mixed B6/129Sv background. Animals and embryos were genotyped by Southern hybridization and/or PCR. Skeletal preparations were made by standard methods²⁶ and *in situ* hybridization was done as described²⁸, as was whole-mount immunostaining²⁸ with 2H3 anti-neurofilament antibody¹⁵ (from Developmental Studies Hybridoma Bank). Detailed protocols and information relating to probes and primers used are available from R.L.J.

Received 9 March; accepted 17 June 1998.

- Cohen, B. *et al.* Fringe boundaries coincide with Notch-dependent patterning centres in mammals and alter Notch-dependent development in *Drosophila*. *Nature Genet.* **16**, 283–288 (1997).
- Johnston, S. H. *et al.* A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development* **124**, 2245–2254 (1997).
- Irvine, K. D. & Wieschaus, E. *fringe*, a boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during *Drosophila* wing development. *Cell* **79**, 595–606 (1994).
- Laufer, E. *et al.* Expression of *Radical fringe* in limb-bud ectoderm regulates apical ectodermal ridge formation. *Nature* **386**, 366–373 (1997).
- Wu, J. Y., Wen, L., Zhang, W. J. & Rao, Y. The secreted product of *Xenopus* gene *lunatic fringe*, a vertebrate signaling molecule. *Science* **273**, 355–358 (1996).
- Panin, V. M., Papayannopoulos, V., Wilson, R. & Irvine, K. D. Fringe modulates Notch–ligand interactions. *Nature* **387**, 908–912 (1997).
- Gossler, A. & Hrabe de Angelis, M. Somitogenesis. *Curr. Top. Dev. Biol.* **38**, 225–287 (1998).
- Christ, B., Schmidt, C., Huang, R., Wilting, J. & Brand-Saberi, B. Segmentation of the vertebrate body. *Anat. Embryol. (Berlin)* **197**, 1–8 (1998).
- Zhang, N. & Gridley, T. Defects in somite formation in *lunatic fringe*-deficient mice. *Nature* **394**, 374–377 (1998).
- Deutsch, U., Dressler, G. R. & Gruss, P. Pax 1, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. *Cell* **53**, 617–625 (1988).
- Neubuser, A., Koseki, H. & Balling, R. Characterization and developmental expression of Pax9, a paired-box-containing gene related to Pax1. *Dev. Biol.* **170**, 701–716 (1995).
- Edmondson, D. G. & Olson, E. N. A gene with homology of the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev.* **3**, 628–640 (1989).
- Mansouri, A. Paired-related murine homeobox gene expressed in the developing sclerotome, kidney, and nervous system. *Dev. Dyn.* **210**, 53–65 (1997).

14. Bettenhausen, B., Hrade de Angelis, M., Simon, D., Guenet, J. L. & Gossler, A. Transient and restricted expression during mouse embryogenesis of *Dll1*, a murine gene closely related to *Drosophila* Delta. *Development* **121**, 2407–2418 (1995).
15. Dodd, J., Morton, S. B., Karagogeos, D., Yamamoto, M. & Jessell, T. M. Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. *Neuron* **1**, 105–115 (1988).
16. Stern, C. D. & Keynes, R. J. Interactions between somite cells: the formation and maintenance of segment boundaries in the chick embryo. *Development* **99**, 261–272 (1987).
17. Takebayashi, K., Akazawa, C., Nakanishi, S. & Kageyama, R. Structure and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-5. Identification of the neural precursor cell-specific promoter element. *J. Biol. Chem.* **270**, 1342–1349 (1995).
18. de la Pompa, J. L. *et al.* Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* **124**, 1139–1148 (1997).
19. Swiatek, P. J., Lindsell, C. E., del Amo, F. E., Weinmaster, G. & Gridley, T. Notch1 is essential for postimplantation development in mice. *Genes Dev.* **8**, 707–719 (1994).
20. Conlon, R. A., Reaume, A. G. & Rossant, J. Notch1 is required for the coordinate segmentation of somites. *Development* **121**, 1533–1545 (1995).
21. Hrade de Angelis, M., McIntyre, J. N. & Gossler, A. Maintenance of somite borders in mice requires the Delta homologue *Dll1*. *Nature* **386**, 717–721 (1997).
22. Williams, R., Lendahl, U. & Lardelli, M. Complementary and combinatorial patterns of Notch gene family expression during early mouse development. *Mech. Dev.* **53**, 357–368 (1995).
23. Dunwoodie, S. L., Henrique, D., Harrison, S. M. & Beddington, R. S. Mouse *Dll3*: a novel divergent Delta gene which may complement the function of other Delta homologues during early pattern formation in the mouse embryo. *Development* **124**, 3065–3076 (1997).
24. Mitsiadis, T. A., Henrique, D., Thesleff, I. & Lendahl, U. Mouse *Serrate-1* (*Jagged-1*): expression in the developing tooth is regulated by epithelial–mesenchymal interactions and fibroblast growth factor-4. *Development* **124**, 1473–1483 (1997).
25. Jen, W. C., Wettstein, D., Turner, D., Chitnis, A. & Kintner, C. The Notch ligand *X-delta-2*, mediates segmentation of the paraxial mesoderm in *Xenopus* embryos. *Development* **124**, 1169–1178 (1997).
26. Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401–1416 (1993).
27. McMahon, A. P. & Bradley, A. The *Wnt-1* (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**, 1073–1085 (1990).
28. Hogan, B., Beddington, R., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Press, New York, 1994).

Acknowledgements. This work was supported by grants from the German National Merit Foundation (to A.A.), the M.D. Anderson Cancer Center Core, March of Dimes, and the NIH (to R.L.J.). We thank R. Beddington, D. Conner, J. Flanagan, A. Gossler, P. Gruss, F. Guillemot, P. Hastay, D. Henrique, B. Hogan, W. Klein, E. Laufer, Y. Mishina, T. A. Mitsiadis and Y. Saga for reagents, T. Gridley for sharing unpublished results, and R. Behringer and C. Tabin for critical comments on the manuscript.

Correspondence and requests for materials should be addressed to R.L.J. (e-mail: rjohnson@odin.mdacc.tmc.edu).

A neuronal ryanodine receptor mediates light-induced phase delays of the circadian clock

Jian M. Ding, Gordon F. Buchanan, Shelley A. Tischkau, Dong Chen, Liana Kuriashkina, Lia E. Faïman, Joan M. Alster^{*}, Peter S. McPherson[†], Kevin P. Campbell[‡] & Martha U. Gillette

Departments of Cell and Structural Biology, Molecular and Integrative Physiology, the Neuroscience Program and * the Computing and Communications Services Office, University of Illinois, Urbana, Illinois 61801, USA

[†] Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 2B4, Canada

[‡] Howard Hughes Medical Institute and Department of Physiology, College of Medicine, University of Iowa, Iowa City, Iowa, 52242, USA

Circadian clocks are complex biochemical systems that cycle with a period of approximately 24 hours. They integrate temporal information regarding phasing of the solar cycle, and adjust their phase so as to synchronize an organism's internal state to the local environmental day and night^{1,2}. Nocturnal light is the dominant regulator of this entrainment. In mammals, information about nocturnal light is transmitted by glutamate released from retinal projections to the circadian clock in the suprachiasmatic nucleus of the hypothalamus. Clock resetting requires the activation of ionotropic glutamate receptors, which mediate Ca^{2+} influx³. The response induced by such activation depends on the clock's temporal state: during early night it delays the clock phase, whereas in late night the clock phase is advanced. To investigate this differential response, we sought signalling elements that contribute solely to phase delay. We analysed intracellular calcium-channel ryanodine receptors, which mediate coupled Ca^{2+} signal-

ling. Depletion of intracellular Ca^{2+} stores during early night blocked the effects of glutamate. Activators of ryanodine receptors induced phase resetting only in early night; inhibitors selectively blocked delays induced by light and glutamate. These findings implicate the release of intracellular Ca^{2+} through ryanodine receptors in the light-induced phase delay of the circadian clock restricted to the early night.

In order to understand the mechanisms underlying directionality in the biphasic light response, we studied the rat suprachiasmatic nucleus (SCN) in a hypothalamic brain slice. The circadian clock persists stably for several days in this preparation⁴. The mean firing frequency of the ensemble of SCN neurons exhibits a spontaneous 24-h oscillation *in vitro* (Fig. 1a). Clock phase was determined from the time-of-peak of this neuronal circadian rhythm. We have previously shown that direct application of Glu to this SCN brain slice induces phase-dependent, light-like clock resetting³. We used

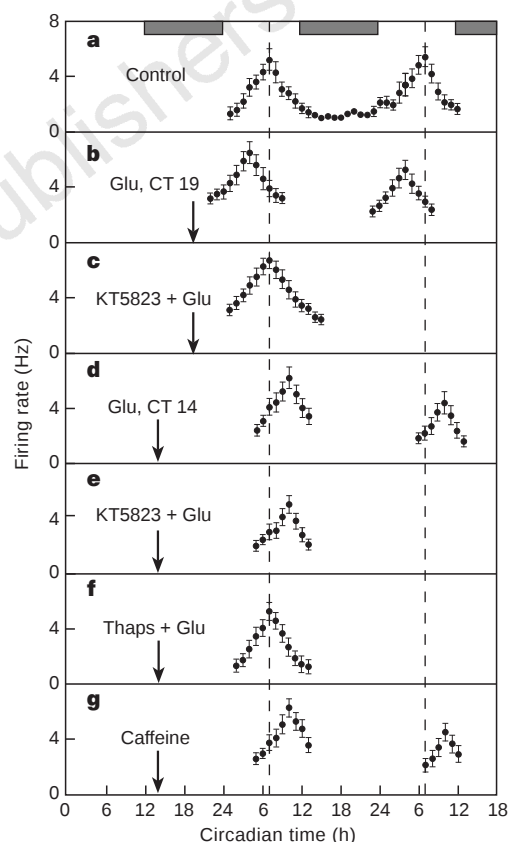


Figure 1 Glu-induced phase advance of SCN neuronal activity rhythm is dependent on PKG, whereas phase delay is not. **a**, The unperturbed SCN neuronal activity over 36 h. The firing rate of this circadian rhythm peaked in mid-subjective day at CT 7, on both day 2 and day 3 *in vitro* in a continuous recording from two SCN. **b**, A microdrop of 10 mM Glu applied at CT 20 advanced the peak of the SCN activity by 3 h ($n = 5$). Neuronal activity was recorded over 10–11 h on each of the next two cycles to define the time-of-peak activity. **c**, Preincubation in 0.1 μM KT5823, a specific PKG inhibitor, blocked the Glu-induced phase advance at CT 20 ($n = 4$, $P < 0.0001$). **d**, At CT 14, Glu delayed the peak of the SCN activity rhythm by 3 h ($n = 6$). **e**, KT5823 did not affect Glu-induced phase delay at CT 14 ($n = 4$, $P > 0.2$). **f**, Preincubation with 0.5 μM thapsigargin, a specific inhibitor of the intracellular Ca^{2+} -ATPase, fully blocked Glu-induced phase delay at CT 14 ($n = 4$, $P < 0.0001$). **g**, A microdrop of 1 mM caffeine to the SCN at CT 14 produced a 3-h phase delay ($n = 4$). Shaded horizontal bars indicate the subjective night of the circadian cycle. The broken vertical lines mark the time of the normal peak of the circadian rhythm of the neuronal activity in unperturbed and EBSS-treated controls. Arrows indicate the time of drug treatment. Statistical analyses applied to this and other microdrop experiments were analyses of variance (ANOVA) with Tukey *post hoc* test.

this finding as a means of studying signalling elements downstream of Glu.

Light-induced clock resetting involves the sequential activation of Glu receptors, Ca^{2+} influx, nitric oxide synthase and intercellular movement of nitric oxide. Nitric oxide can activate soluble guanylyl cyclase, which increases cGMP and activates cGMP-dependent protein kinase (PKG). Activation of cGMP-dependent pathways of SCN in brain slices has been shown to stimulate phase advance of the clock at night, but not during the day^{5,6}. Furthermore, intracerebroventricular (i.c.v.) injection of KT5823, a specific PKG inhibitor, blocks light-induced advances of wheel-running rhythms in hamsters during the late night^{7,8}. To determine whether PKG activation is associated with the Glu-induced phase shifts, SCN slices were bathed in KT5823 before treatment with Glu or media. A 30-min incubation in 0.1 μM KT5823 fully inhibited PKG phosphotransferase activity in the SCN during stimulation by Glu at circadian time (CT; time after entrained lights on) 14 and 20. KT5823 itself had no effect on clock phase in late or early night, but it completely blocked phase advances of the neuronal activity rhythm induced by Glu in the late night (Fig. 1b, c). However, in

the early night, KT5823 had no effect on Glu-induced phase delays (Fig. 1d, e). Therefore, the NO-GC-cGMP-PKG pathway does not mediate the Glu-induced phase delay.

To investigate alternative pathways, we evaluated intracellular Ca^{2+} (Ca_i^{2+}) signalling. Thapsigargin depletes Ca_i^{2+} by blocking the Ca^{2+} -ATPase that replenishes Ca_i^{2+} stores in the endoplasmic reticulum⁹. When applied at either at CT 20 or 14, thapsigargin itself did not alter the phasing of the neuronal activity rhythm. Thapsigargin treatment had only a small inhibitory effect on the phase advance induced by Glu (-1.15 h, CT 20, $n = 4$, $P < 0.001$), but it completely blocked the Glu-induced phase delay (Fig. 1f). This finding indicates that the Glu-induced phase delay requires Ca_i^{2+} release.

To examine the Ca_i^{2+} signalling mechanism in early night, we probed the Ca_i^{2+} -channel ryanodine receptor (RyR). RyRs are integral membrane proteins associated with cellular organelles, such as the endoplasmic reticulum, that sequester Ca^{2+} . They are expressed in neurons where they can mediate release of stored Ca_i^{2+} (ref. 10). Ca_i^{2+} mobilization by modulators of RyRs can reset circadian rhythms of melatonin production in avian pinealocytes¹¹. A preliminary report¹² suggests that the rat SCN exhibits circadian variation in the binding characteristics of RyRs, whereas inositol-1,4,5-trisphosphate (IP_3) receptor binding remains constant across the 24-h cycle.

To evaluate the contribution of RyRs to phase delays, we studied a range of reagents that have different mechanisms of action but all release Ca_i^{2+} through RyRs. Caffeine can directly activate RyRs^{10,13}, and FK506 and rapamycin bind to immunophilins, a class of proteins that regulate the ion-channel functions of RyRs¹³⁻¹⁵. Although FK506 and rapamycin exert their immunosuppressive actions through different effectors, both bind to FKBP12, a well-studied immunophilin stabilizer of RyRs¹³⁻¹⁵. Application of caffeine, FK506 or rapamycin to the SCN at CT 14 induced Glu-like

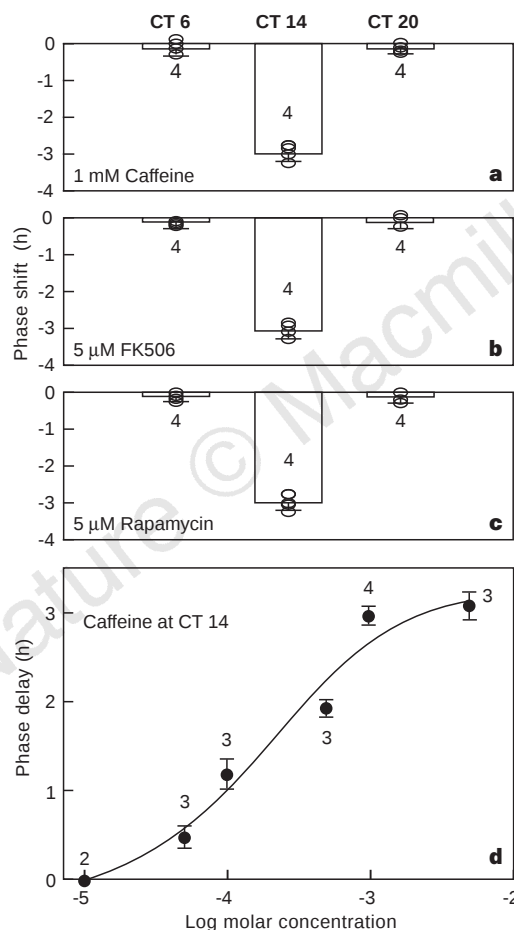


Figure 2 SCN sensitivity to caffeine and to the immunophilin ligands FK506 and rapamycin is restricted to early night. **a-c**, Temporal sensitivity to caffeine (**a**), FK506 (**b**) or rapamycin (**c**) was assessed at mid-day (CT 6), early night (CT 14) and late night (CT 20). These reagents did not affect phasing of the circadian rhythm when applied at CT 6 or CT 20. However, at CT 14, each drug induced a 3-h phase delay. Graph shows data points, mean, s.e.m. and n (number). The circadian time effect was significant for each drug ($P < 0.0001$). The phase shifts at CT 14 were significantly different from those at CT 6 and at CT 20 ($P < 0.0001$, each comparison, each drug). **d**, Dose-response curve of caffeine at CT 14. Half-maximum response occurred with a 0.1 mM microdrop. Each data point represents mean $\pm$ s.e.m. The regression curve was fit to the data points using the Hill equation.

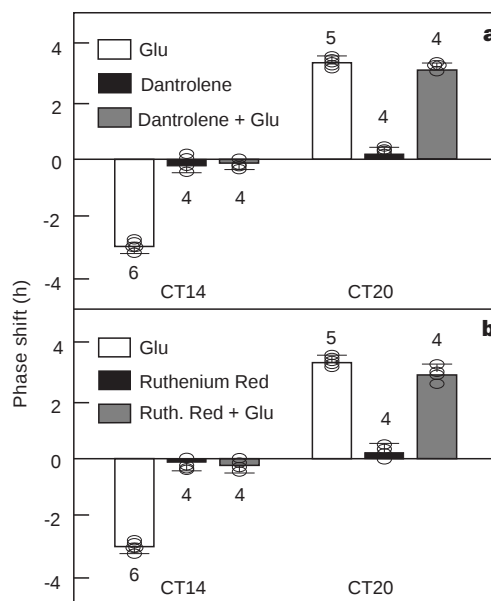


Figure 3 RyR inhibitors selectively block Glu-induced phase delays in the early night. **a, b**, Bath application of 20 μM dantrolene (**a**) or 10 μM ruthenium red (**b**) to the brain slice before stimulation by Glu blocked the phase delay normally induced by Glu at CT 14, but not the advance at CT 20. At CT 14, for both inhibitors, the phase delay induced by Glu was significantly different ($P < 0.0001$) from that for the inhibitor alone or the inhibitor plus Glu ($P < 0.0001$). The shifts for inhibitor and inhibitor plus Glu did not differ significantly from one another. At CT 20, for both inhibitors, the phase advances for Glu and Glu plus inhibitor did not differ significantly from each other, but they each did differ significantly ($P < 0.0001$) from the shifts seen with inhibitor alone. The Glu controls at CT 14 and CT 20 in **a** are replotted in **b**.

phase delays of ~3 h (Fig. 1g); these treatments did not affect clock phase at CT 6 or 20 (Fig. 2a–c). Microdrops containing 10^{-6} to 10^{-2} M caffeine elicited a dose-dependent phase delay, with a half-maximal response near 5×10^{-4} M (Fig. 2d).

To examine further the pathway mediating Glu-induced phase delay, we used inhibitors of RyRs. Dantrolene and ruthenium red each effectively block activation of RyRs, although they act at different sites on the RyR molecule^{13,16}. Preincubating the SCN in either dantrolene or ruthenium red before applying a microdrop of Glu at CT 14 fully blocked phase delays. These inhibitors had no effect on Glu-induced phase advances at CT 20 (Fig. 3). To determine whether RyR inhibition affects light-stimulated phase shifts, we tested dantrolene *in vivo*. When photic stimuli were evaluated after i.c.v. injection of dantrolene, the phase delays of the wheel-running rhythm of hamsters under constant darkness were significantly attenuated (Fig. 4).

Because caffeine and FK506 have diverse cellular actions^{17,18}, we

sought to identify the pathway mediating the effects of FK506 and caffeine on the phase-delaying process. The efficacy of these reagents at CT 14 was tested against the RyR antagonist dantrolene. In each case, dantrolene fully blocked the agonist's effect ($n = 4$, $P < 0.0001$, each condition; data not shown). Western blot analysis with affinity-purified anti-neuronal RyR antibody^{19,20} demonstrated that the SCN expresses neuronal RyR (relative molecular mass 325K, CT 14, $n = 3$; data not shown). Because dantrolene is highly specific for RyRs and antagonizes the effects of these modulators, FK506 and caffeine are probably affecting phase change within the SCN through their RyR-binding properties.

Our data show that SCN signalling elements diverge downstream from Glu. In the late night, the light/Glu signal causes a 3-h phase advance of the timekeeping mechanism by means of a cGMP-dependent pathway (Fig. 1b, c)^{7,8,21}. In the early night, however, signal transduction downstream of Glu follows a distinct pathway leading to activation of RyRs and release of Ca^{2+} , resulting in a phase delay of 3 h. Although the steps linking Glu and RyRs are unknown, nitric oxide is a likely intermediary that could stimulate the ADP ribosyl cyclase pathway^{22,23} and/or directly activate RyRs by means of nitric oxide-mediated polynitrosylation²⁴. Further, our findings demonstrate that RyR activation contributes a pivotal directional signal in the context of dynamic clock state. Although their roles in neurons are not fully understood, RyRs are thought to modulate the duration and/or amplitude of the Ca^{2+} signal^{25,26}. Potential mechanisms by which an increase in Ca^{2+} could contribute to the phase-delaying process include activation of Ca^{2+} -dependent kinases, phosphatases and/or proteases leading to the inactivation or degradation of clock regulatory elements. Indeed, light-induced protein degradation has been reported in the regulation of clock-related genes, such as *tim* in *Drosophila*²⁷. □

Methods

Electrophysiology and pharmacology. Inbred Long–Evans rats 6–10 weeks old were used in these experiments. Procedures for brain slice preparation and electrophysiology have been described previously^{3,4}. Briefly, spontaneous activity of single SCN neurons was sampled extracellularly for 4-min periods. From sequential unit activities, 2-h running means of the neuronal ensemble were calculated. The phase of the underlying circadian clock was determined by the time of the peak in the oscillation^{3,4}. Under constant conditions *in vitro*, the unperturbed sinusoidal pattern of neuronal activity is predictably high during the subjective day and low during the subjective night. Activity peaks mid-subjective day, at circadian time 7 (CT 7, 7 h after lights on in the rat colony; 12:12 LD). To evaluate experimental stimuli, the perfusion pump was stopped, a droplet of 0.2 or 0.5 μl of test substance was applied bilaterally to the SCN for 10 min, and the SCN was then rinsed with glucose- and bicarbonate-supplemented Earle's balanced salt solution (EBSS). To evaluate potential inhibitors, perfusion medium was replaced with medium containing the antagonist for 20 min before microdrop application of the phase-shifting stimulus. To determine the phase after drug treatments, the time-of-peak neuronal activity was assessed for 1–2 days.

Stereotactic surgery for i.c.v. cannulation and drug injection. Male Syrian hamsters (Harlan, 100 g) were used to assess pharmacological effects on behavioural rhythms^{7,8}. A 23-gauge guide cannula was implanted to the lateral ventricle of each (0.4 mm rostral to bregma, 3.0 mm ventral to dura, 2.4 mm lateral to midline, with bregma and lambda in a levelled plain). A 30-gauge stylet was placed in the guide cannula to maintain patency. For i.c.v. injections the stylet was removed and a 30-gauge injector attached to a microsyringe was inserted. At 20 min before light exposure, the animals were lightly anaesthetized with methoxyflurane during i.c.v. injection. Each animal received a 1- μl injection under dim red light (<1 lux) over 30 s. Generally, animals began to wake up before completion of the injection and were completely awake before light exposure (500 lux, 15 min). The animals were then returned to their cages and maintained in darkness for 10–14 days before the next treatment.

Received 11 November 1997; accepted 18 May 1998.

1. Pittendrigh, C. S. Circadian rhythms and the circadian organization of living organisms. *Cold Spring Harb. Symp. Quant. Biol.* **25**, 159–184 (1960).

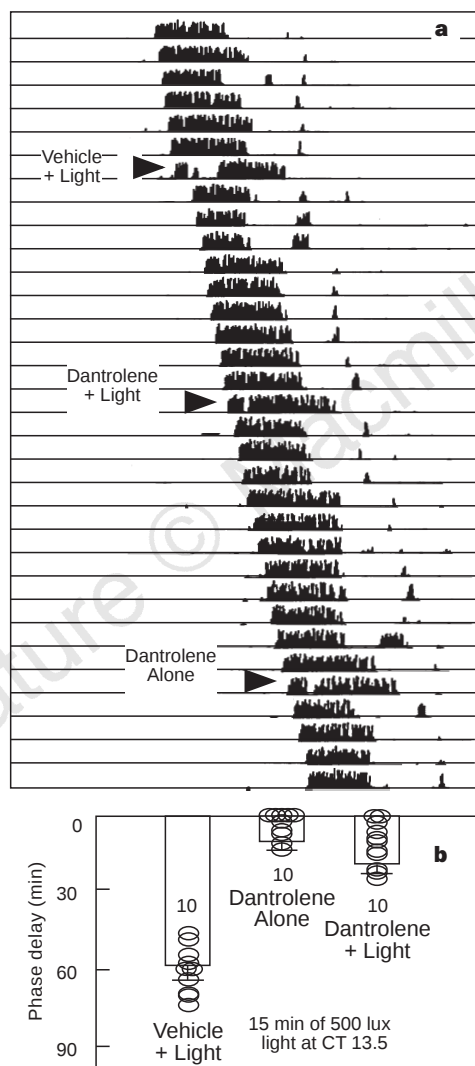


Figure 4 Dantrolene blocks light-induced phase delays of wheel-running activity rhythm of free-running hamsters. **a**, The actogram of one hamster throughout the various treatments. Each horizontal line represents 24 h. Light exposure at CT 13.5 after vehicle (1 μl) administered i.c.v. 20 min earlier induced a 60-min phase delay. Dantrolene (50 μM , 1 μl , i.c.v.) administered before light exposure significantly attenuated the light-induced phase delay. Dantrolene without light did not affect the phase. **b**, The phase delay for the vehicle plus light group was significantly different from that for either of the dantrolene-treated groups ($P < 0.0001$, Kruskal–Wallis rank ANOVA), but the dantrolene-treated groups did not differ significantly from each other.

2. De Coursey, P. J. Daily light sensitivity rhythm in a rodent. *Science* **131**, 33–35 (1960).
3. Ding, J. M. *et al.* Resetting the biological clock: mediation of nocturnal circadian shifts by glutamate and NO. *Science* **266**, 1713–1717 (1994).
4. Prosser, R. A. & Gillette, M. U. The mammalian circadian clock in the suprachiasmatic nucleus is reset *in vitro* by cAMP. *J. Neurosci.* **9**, 1073–1081 (1989).
5. Prosser, R. A., McArthur, A. J. & Gillette, M. U. cGMP induces phase shifts of a mammalian circadian pacemaker at night, in antiphase to cAMP effects. *Proc. Natl Acad. Sci. USA* **86**, 6812–6815 (1989).
6. Liu, C., Ding, J. M., Faiman, L. E. & Gillette, M. U. Coupling of muscarinic cholinergic receptors and cGMP in nocturnal regulation of the suprachiasmatic circadian clock. *J. Neurosci.* **17**, 659–666 (1997).
7. Weber, E. T., Gannon, R. L. & Rea, M. A. cGMP-dependent protein kinase inhibitor blocks light-induced phase advances of circadian rhythms *in vivo*. *Neurosci. Lett.* **197**, 227–230 (1995).
8. Mathur, A., Golombek, D. A. & Ralph, M. R. cGMP-dependent kinase inhibitors block light-induced phase advances of circadian rhythms *in vivo*. *Am. J. Physiol.* **270**, R1031–R1036 (1996).
9. Thastrup, O., Cullen, P. J., Drobak, B. K., Hanley, M. R. & Dawson, A. P. Thapsigargin, a tumor promoter, discharges Ca^{2+} stores by specific inhibition of the endoplasmic reticulum Ca^{2+} -ATPase. *Proc. Natl Acad. Sci. USA* **87**, 2466–2470 (1990).
10. McPherson, P. S. *et al.* The brain ryanodine receptor: A caffeine-sensitive calcium release channel. *Neuron* **7**, 17–25 (1991).
11. Zatz, M. & Heath, J. R. Calcium and photorefractive in chick pineal cells revisited: effects of caffeine, thapsigargin, EGTA, and light on the melatonin rhythm. *J. Neurochem.* **65**, 1332–1341 (1995).
12. Dent, M., Diaz-Munoz, M., Chavez, J. L. & Aguilar-Roblero, R. Circadian variations of ryanodine receptor in the suprachiasmatic nucleus of the rat. *Neurosci. Abstr.* **22**, 139.9 (1996).
13. Brillantes, A. B. *et al.* Stabilization of calcium release channel (ryanodine receptor) function by FK506-binding protein. *Cell* **77**, 513–523 (1994).
14. Kaftan, E., Marks, A. R. & Ehrlich, B. E. Effects of rapamycin on ryanodine receptor/ Ca^{2+} -release channels from cardiac muscle. *Circ. Res.* **78**, 990–997 (1996).
15. Snyder, S. H. & Sabatini, D. M. Immunophilins and the nervous system. *Nature Med.* **1**, 32–37 (1995).
16. Parness, J. & Palnikar, S. S. Identification of dantrolene binding sites in porcine skeletal muscle sarcoplasmic reticulum. *J. Biol. Chem.* **270**, 18465–18472 (1995).
17. Fredholm, B. B. Adenosine, adenosine receptors and the actions of caffeine. *Pharmacol. Toxicol.* **76**, 93–101 (1995).
18. Kunz, J. & Hall, M. N. Cyclosporin A, FK506 and rapamycin: more than just immunosuppression. *Trends Biochem. Sci.* **18**, 334–338 (1993).
19. McPherson, P. S. & Campbell, K. P. Solubilization and biochemical characterization of the high affinity ryanodine receptor from rabbit brain membranes. *J. Biol. Chem.* **265**, 18454–18460 (1990).
20. McPherson, P. S. & Campbell, K. P. Characterization of the major brain form of the ryanodine receptor/ Ca^{2+} release channel. *J. Biol. Chem.* **268**, 19785–19790 (1993).
21. Weber, E. T., Gannon, R. L., Michel, A. M., Gillette, M. U. & Rea, M. A. Nitric oxide synthase inhibitor blocks light-induced phase shifts of the circadian activity rhythm, but not *c-fos* expression in the suprachiasmatic nucleus of the Syrian hamster. *Brain Res.* **692**, 137–142 (1995).
22. Sitsapesan, R., McGarry, S. J. & Williams, A. J. Cyclic ADP-ribose, the ryanodine receptor and Ca^{2+} release. *Trends Pharmacol. Sci.* **16**, 386–391 (1995).
23. Pawlikowska, L., Cottrell, S. E., Harms, M. B., Li, Y. & Rosenberg, P. A. Extracellular synthesis of cADP-ribose from nicotinamide-adenine dinucleotide by rat cortical astrocytes in culture. *J. Neurosci.* **16**, 5372–5881 (1996).
24. Xu, L., Eu, J. P., Meissner, G. & Stamler, J. S. Activation of the cardiac calcium release channel (ryanodine receptor) by poly-nitrosylation. *Science* **279**, 234–237 (1998).
25. Obenaus, A., Mody, I. & Baimbridge, K. G. Dantrolene- Na^{+} blocks induction of long-term potentiation in hippocampal slices. *Neurosci. Lett.* **98**, 172–178 (1989).
26. O'Mara, S. M., Rowan, M. J. & Anwyl, R. Dantrolene inhibits long-term depression and depotentiation of synaptic transmission in the rat dentate gyrus. *Neuroscience* **68**, 621–624 (1995).
27. Hunter-Ensor, M., Ousley, A. & Sehgal, A. Regulation of the *Drosophila* protein *timeless* suggests a mechanism for resetting the circadian clock by light. *Cell* **84**, 677–685 (1996).

Acknowledgements. We thank P. Best and R. Gillette for critical evaluation of the manuscript; T. Weber and C. Colwell for technical advice on the i.c.v. procedure; P. Imrey for statistical advice. These studies were supported by grants from the NINDS of the NIH and a university scholar award, UIUC (M.U.G.).

Correspondence and requests for materials should be addressed to M.U.G. (e-mail: mgillett@uiuc.edu).

High-frequency firing helps replenish the readily releasable pool of synaptic vesicles

Lu-Yang Wang[†] & Leonard K. Kaczmarek[†]

^{*} Division of Neurology & The Program in Brain and Behavior, The Hospital for Sick Children and Department of Physiology, University of Toronto, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada

[†] Department of Pharmacology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, USA

Synapses in the central nervous system undergo various short- and long-term changes in their strength^{1–3}, but it is often difficult to distinguish whether presynaptic or postsynaptic mechanisms are responsible for these changes. Using patch-clamp recording from giant synapses in the mouse auditory brainstem^{4–7}, we show here that short-term synaptic depression can be largely attributed to rapid depletion of a readily releasable pool of vesicles. Replenishment of this pool is highly dependent on the recent history of

synaptic activity. High-frequency stimulation of presynaptic terminals significantly enhances the rate of replenishment. Broadening the presynaptic action potential with the potassium-channel blocker tetraethylammonium, which increases Ca^{2+} entry, further enhances the rate of replenishment. As this increase can be suppressed by the Ca^{2+} -channel blocker Cd^{2+} or by the Ca^{2+} buffer EGTA, we conclude that Ca^{2+} influx through voltage-gated Ca^{2+} channels is the key signal that dynamically regulates the refilling of the releasable pool of synaptic vesicles in response to different patterns of inputs.

Glutamatergic excitatory postsynaptic currents (EPSCs) were recorded from the principal neurons of the medial nucleus of the trapezoid body (MNTB) in an all-or-none fashion, consistent with the fact that each postsynaptic neuron is innervated at the soma by a single presynaptic calyx (the calyx of Held)^{8,9}. When the presynaptic axon was stimulated with a 100-ms train at a frequency of 100 to 300 Hz, we observed a frequency-dependent depression in the amplitude of the EPSCs (Fig. 1a, b). This depression was short-lasting and reversed within 15 s.

To investigate the mechanisms causing this short-term depression, we first tested whether desensitization of postsynaptic AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptors caused a decline in the amplitude of the EPSCs. When the desensitization was blocked with cyclothiazide (CTZ, 0.1 mM)^{10,11}, the decay time course of individual EPSCs was prolonged, but the relative amplitude of the last EPSC in the train was about the same as that in the absence of CTZ at all frequencies. A typical recording at 200 Hz is shown in Fig. 1c, d. Because these experiments were carried out under conditions of a high quantal output (in 2 mM

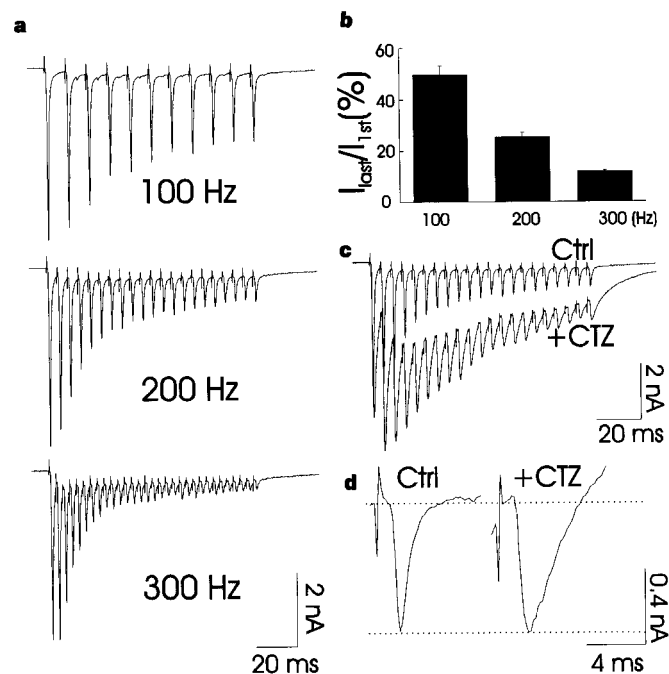


Figure 1 Use-dependent synaptic depression is independent of postsynaptic desensitization. **a**, Typical recording showing frequency-dependent depression in EPSCs in response to a 100-ms train of stimulation at 100, 200 or 300 Hz. Each EPSC is preceded by a stimulation artefact. The holding potential was -60 mV for this and subsequent figures unless otherwise indicated. **b**, Summary of the extent of synaptic depression at three frequencies. On average, the amplitude of the last EPSC in the train, at a frequency of 100, 200 and 300 Hz, was $49.7 \pm 3.5\%$, $25.3 \pm 1.8\%$ and $11.8 \pm 0.5\%$ of that of the first EPSC ($n = 10$), respectively. **c**, EPSCs recorded from the same cell as in **a** before and after addition of CTZ (0.1 mM) are superimposed to show that blockade of AMPA-receptor desensitization failed to eliminate synaptic depression. **d**, Comparison of the last EPSC with or without CTZ treatment from **a**, after adjusting the baseline to the same level.

extracellular Ca^{2+} ($[\text{Ca}^{2+}]_o$), in which any presynaptic effects of CTZ^{11,12} are probably minimized, we conclude that postsynaptic desensitization does not account for the synaptic depression and that the reduction in EPSCs is likely to reflect a presynaptic change.

An alternative explanation for the decline in EPSCs is that there may be a use-dependent decrease in presynaptic Ca^{2+} current during repetitive action-potential firing. To test this, we recorded Ca^{2+} current directly from presynaptic calyces. The Ca^{2+} current was activated by depolarizing steps positive to -40 mV and reached maximal activation at -10 mV (Fig. 2a, b). Addition of $50 \mu\text{M}$ cadmium, a non-selective Ca^{2+} channel blocker, completely eliminated this current (Fig. 2b). In response to a sustained depolarizing step for 100 ms (from -70 to -10 mV) we observed 32% inactivation ($n = 4$) (Fig. 2d). In contrast, Ca^{2+} tail currents repetitively evoked at 100 to 300 Hz showed little reduction over the 100-ms test period and a slight facilitation often preceded the decline in these currents (Fig. 2c, d). The difference in the time course of synaptic depression and the very small reduction in Ca^{2+} currents suggests that the synaptic depression is unlikely to be attributed to inactivation or feedback inhibition of Ca^{2+} currents by activating presynaptic metabotropic glutamate receptors⁵. Furthermore, we observed a marked reduction in the synaptic depression or even a facilitation in EPSCs, when extracellular Ca^{2+} was lowered from 2 mM to 1 or 0.5 mM¹³, suggesting that the depression is likely to be a result of reducing the size of the readily releasable pool of synaptic vesicles or even depleting it at high frequencies (at 300 Hz for example), as suggested for other synapses^{14–19} and for the same synapse in rats²⁰. However, we do not exclude the possibility that the build-up of an unknown inhibitory factor or a change in the calcium sensitivity of the release machinery may also contribute.

The calyx of Held synapse preserves timing information for

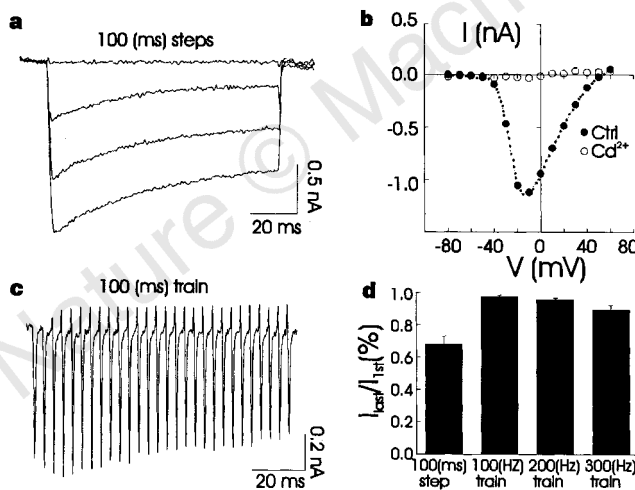


Figure 2 Presynaptic Ca^{2+} currents do not inactivate significantly in response to repetitive stimulation. **a**, Ca^{2+} currents recorded from a presynaptic calyx by depolarizing steps (100 ms) from -70 mV to -50 , -10 , $+10$ and $+30$ mV. The holding potential was -70 mV and leak currents were subtracted on-line with the P/4 protocol. **b**, Current-voltage (I - V) relationships of the Ca^{2+} current from **a** in the absence (black circles) and presence (white circles) of $50 \mu\text{M}$ Cd^{2+} . The measurement of current amplitude was made within the last 5 ms of the voltage step. **c**, Ca^{2+} tail currents repetitively evoked by a train of voltage commands with the waveform of action potentials (half-width, 0.6 ms; height, 130 mV) from -70 mV to $+50$ mV at 300 Hz. The capacitive transients and passive currents were measured in the presence of $50 \mu\text{M}$ Cd^{2+} and subtracted from the control. **d**, Plot summarizing the extent of Ca^{2+} current inactivation in response to a sustained step (-70 to -10 mV, 100 ms) or repetitive short steps (0.9 ms) at 100 and 200 Hz ($n = 4$) or action potential waveform commands at 300 Hz ($n = 3$). The Ca^{2+} current amplitude at the end of the 100 ms voltage step is $67.9 \pm 5.0\%$ of its peak amplitude, while the last Ca^{2+} tail current amplitude at 100, 200 and 300 Hz is $97.4 \pm 0.6\%$, $95.6 \pm 0.8\%$ and $89.1 \pm 3.0\%$ of the first tail current, respectively.

sound localization in the auditory pathway^{21–23}. To ensure synaptic fidelity, particularly at high frequencies, the readily releasable pool has to be refilled rapidly following depletion. To determine the time course of replenishment, we first delivered a short train (100 ms) of stimuli to the presynaptic terminal and then gave a single test stimulus at different intervals following each train (Fig. 3). We found that the time course of refilling the pool following 100-Hz stimulation was well described by a single exponential with a time constant of 5 s ($n = 6$) (Fig. 3a, c). However, when the presynaptic terminal was stimulated at 300 Hz, there was a rapid refilling phase within the first 300 ms, which surpassed the recovery level seen at 100 Hz (Fig. 3b, c). This rapid phase was followed by a slower replenishment similar to that at 100 Hz. The overall time course at 300 Hz was fitted by a double-exponential function with time

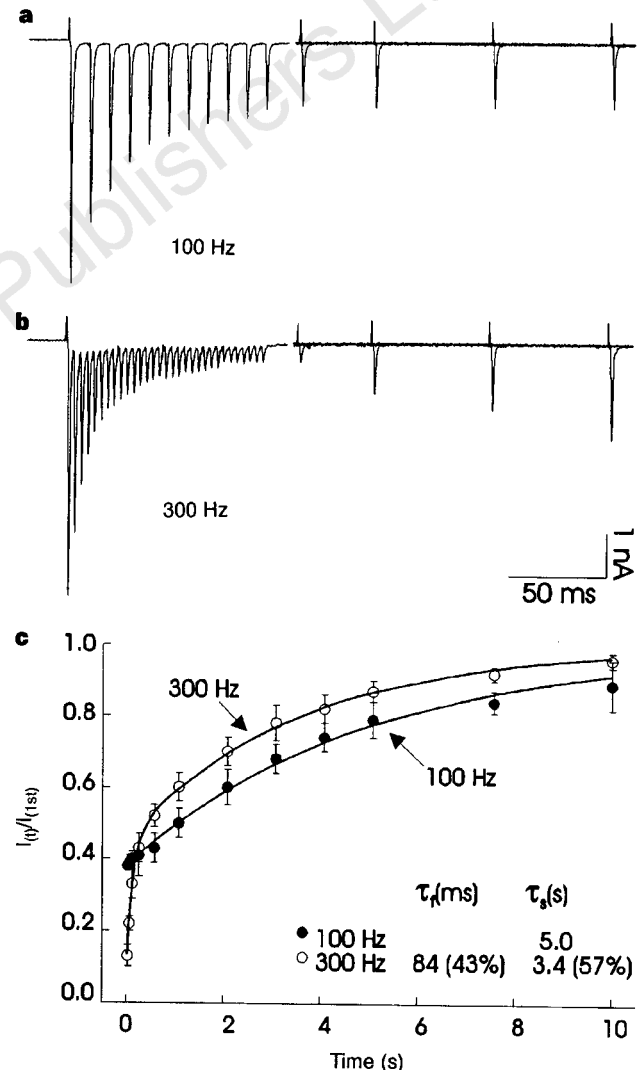


Figure 3 Replenishment kinetics are dependent on the frequency of presynaptic inputs. **a**, **b**, Recording from the same synapse showing different refilling kinetics following a 100-ms train of stimuli at 100 and 300 Hz. Following each test train, a single EPSC at different time intervals from the last EPSC of the train was evoked to measure the extent of refilling. Traces recorded during the test trains are averaged and shown as a single sweep. Stimulation artefacts in the test train are removed for clarity. The EPSCs during recovery are superimposed. The interval between the test train sweeps was 20 s. **c**, Averaged time course at 100 and 300 Hz from six synapses. Note the appearance of an early phase of recovery at 300 Hz. In this and subsequent figures, unless otherwise indicated, smooth curves are single or double exponential fit to the averaged data, and τ_f and τ_s are time constants for fast and slow refilling phases, respectively. Percentages of contributions of fast and slow components are given in parentheses.

constants of 84 ms and 3.4 s, respectively ($n = 6$) (Fig. 4c). Thus, the replenishment kinetics appear to be strongly influenced by preceding synaptic activity at this synapse. A similar frequency-dependent recovery from depression in EPSPs has been noted at a locust motor neuron synapse²⁴.

Because direct recordings from calyces showed that the waveform of presynaptic action potentials at 100 or 300 Hz showed little change during the test train (Fig. 4a, b) or recovery (Fig. 4a, c), our results indicate that the pattern of synaptic activity may affect the rate of replenishing the readily releasable pool. We considered that Ca^{2+} influx through voltage-gated Ca^{2+} channels during a train of action potentials not only triggers synaptic release but also regulates the refilling kinetics of the readily releasable pool. To test this idea, we broadened the width of presynaptic action potentials using the potassium-channel blocker tetraethylammonium (TEA; 1 mM) to introduce more Ca^{2+} into the terminal during high-frequency firing. TEA selectively blocked a high-threshold potassium current (data not shown) and broadened the width of action potentials, with little effect on their height (Fig. 4d–f). It did not reduce the ability of presynaptic terminals to fire action potentials at frequencies of up to 200 Hz (Fig. 4d). Again, the spike waveform during the test train or recovery was very similar, except that a slight broadening near the end of the train was induced in the presence of TEA

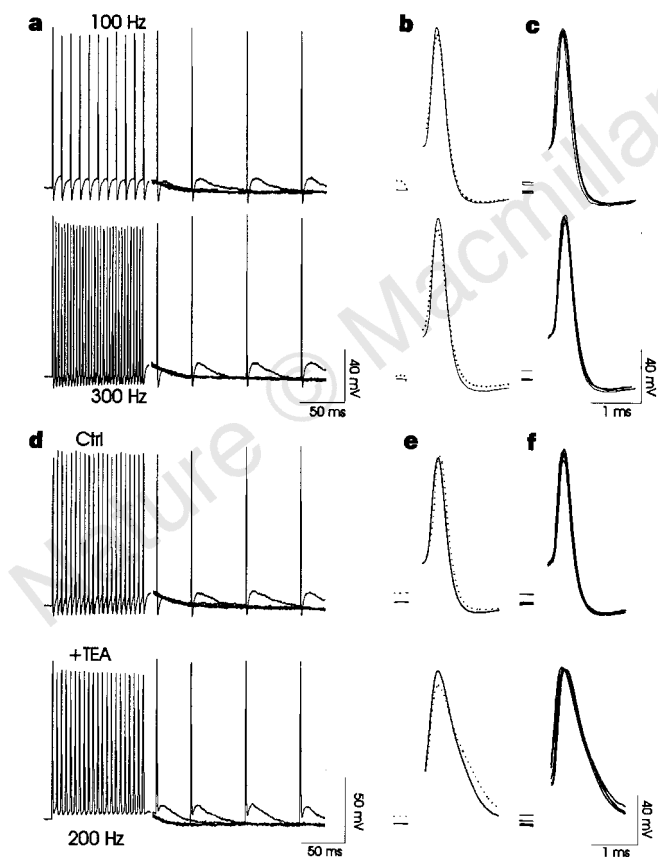


Figure 4 Presynaptic spike waveform and TEA-induced spike broadening. **a**, Train of action potentials evoked at 100 (top) or 300 (bottom) Hz was followed by four test spikes at the same time points as those used for measuring recovery from synaptic depression. **b**, The first spike (solid line) and the last spike (dotted line) within the test train are superimposed. **c**, First spike of the train and all four test spikes following the train are superimposed. **d–f**, Same experiment as in **a** before (top) and after (bottom) application of TEA (1 mM) is shown, except that the frequency of the test train was 200 Hz. Note that there is a slight broadening and reduction in amplitude of the last spike of the test train in the presence of TEA (dotted line in **e**). Stimulation artefacts preceding the upstroke of spikes in **b**, **c**, **e** and **f** are removed for clarity. A depolarizing after-potential following individual spikes was often observed in presynaptic recordings⁷.

(Fig. 4e, lower panel).

Application of 1 mM TEA caused a slight synaptic delay and increased the size of the EPSCs by $65.1 \pm 7.0\%$ ($n = 7$) at a low stimulation frequency of 0.05 Hz (Fig. 5a). Under control conditions, the refilling time course following a 100-ms train at 200 Hz was well described by a single exponential curve with a time constant of 3.8 s (Fig. 5b, f). However, in the presence of TEA there was a striking enhancement in the replenishment kinetics following the test train (Fig. 5c, f) and the refilling curve is clearly biphasic, with a fast time constant of 52 ms and a slow time constant of 1.1 s. Most refilling can be attributed to the early phase that occurs within the first 300 ms after the last stimulus of the train (Fig. 5c, f). To test the hypothesis that Ca^{2+} mediates the rapid phase of replenishment, we applied Cd^{2+} to reduce Ca^{2+} influx through TEA-broadened action potentials. We found that Cd^{2+} , at the low concentration of $1.5 \mu\text{M}$, only reduced a fraction of phasic release ($24.9 \pm 3.9\%$; $n = 7$) (Fig. 5d), but reversed the TEA-induced enhancement of replenishment kinetics, resulting in a single exponential refilling time course with a time constant of 2.2 s (Fig. 5e, f).

To demonstrate further that Ca^{2+} directly mediates the early phase of the replenishment, we introduced a slow Ca^{2+} buffer into presynaptic calyces by preincubating slices with the tetra-acetoxymethyl ester of EGTA (EGTA-AM) (0.2 mM). Such treatment only slightly reduced the extent of synaptic depression (<10%) at all three frequencies (Fig. 6). In contrast, the fast component of the refilling at 300 Hz was completely eliminated

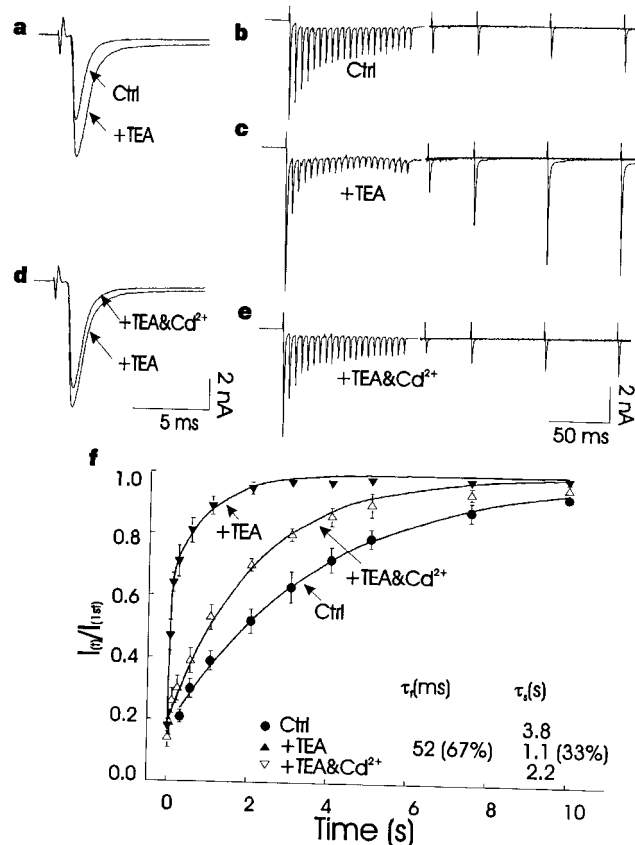


Figure 5 Ca^{2+} enhances the replenishment of the readily releasable pool of synaptic vesicles. **a**, **d**, TEA increased phasic release measured at 0.05 Hz and a fraction of this increase was reversed by co-application of Cd^{2+} . **b**, **c**, **e**, Example of different refilling time courses following a 100-ms train at 200 Hz before and after addition of 1 mM TEA or TEA plus $1.5 \mu\text{M}$ Cd^{2+} in the same synapse as in **a**, **d**, **f**. Pooled data from seven synapses showing that TEA caused a marked shift in the refilling time course and that this shift can be partially reversed by Cd^{2+} . Time constants and percentages of contributions of fast and slow components are shown.

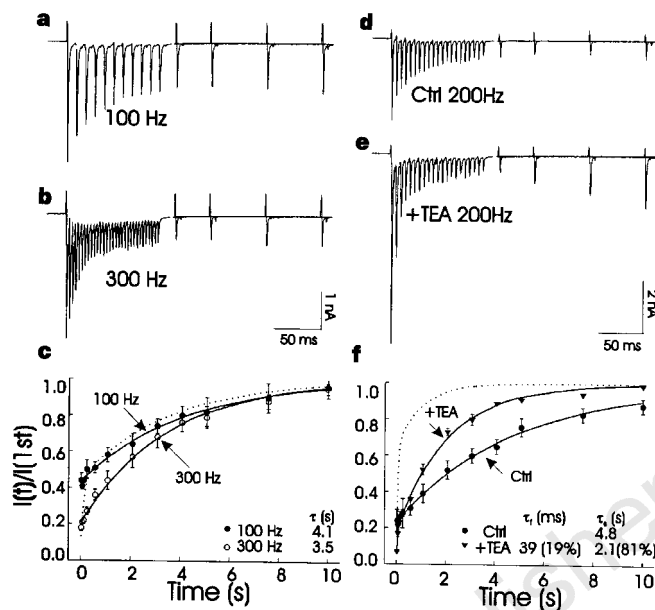


Figure 6 The Ca^{2+} chelator EGTA blocks the fast phase of replenishment. **a, b**, Example showing that EGTA-AM pretreatment blocked the fast component of recovery at 300 Hz. **c**, Plot summarizing refilling time course following the test train at 100 Hz (black circles) and 300 Hz (white circles) from four synapses. **d, e**, Recording from another synapse showing that TEA-induced facilitation in the recovery at

200 Hz became less prominent in EGTA-AM-pretreated synapses. **f**, Summary of experiments done as in **d** and **e**, from four synapses before (circles) and after (triangles) TEA application. Time constants and percentages of contributions of fast and slow components are given. Dotted lines in **c** and **f** are the control recovery time course taken from Figs 3 and 5 and are shown for comparison.

(Fig. 6a–c). Application of TEA in EGTA-AM-treated slices produced a greater enhancement of the EPSC ($105.2 \pm 18.2\%$; $n = 4$) than in untreated slices (65%, Fig. 5a). The TEA-induced early refilling component in control slices was, however, severely attenuated by EGTA-AM pretreatment (Fig. 6d–f). These results show that Ca^{2+} entry into the presynaptic terminal regulates the replenishment of the readily releasable pool of synaptic vesicles.

It is possible that Ca^{2+} influx during action potential firing results in Ca^{2+} -induced release of Ca^{2+} from internal stores, which may in turn affect replenishment^{24,25}. However, a depletion of internal Ca^{2+} stores by pretreating slices with thapsigargin (3 μM), an inhibitor of the Ca^{2+} pump on internal stores, did not block TEA-induced enhancement of refilling (data not shown; $n = 3$), suggesting that Ca^{2+} influx through voltage-gated Ca^{2+} channels can act as a direct signal for the facilitation. An alternative possibility is that the enhanced recovery from synaptic depression in the presence of TEA reflects an acceleration of the removal of postsynaptic desensitization. To address this possibility, we determined the recovery time course in TEA alone, or in combination with Cd^{2+} ($n = 5$) in the presence of CTZ (0.1 mM). We found that CTZ itself only produced a marginal enhancement of the recovery (time constant: 4 s in the control; 2.9 s in CTZ). However, it failed to prevent the appearance of TEA-induced biphasic recovery (time constants: 79 ms and 1.6 s). Cd^{2+} again readily suppressed the action of TEA (time constant, 1.8 s). These experiments indicate that Ca^{2+} -induced Ca^{2+} release from internal stores, or removal of the postsynaptic desensitization of AMPA receptors, does not play a dominant role in regulating the rapid recovery from synaptic depression at high frequencies.

We have demonstrated that depletion of the readily releasable pool of synaptic vesicles largely accounts for the short-term depression at the calyx of Held synapse. The time constant for fully replenishing this pool following a 100-ms train at 100 Hz is 5 s, a value comparable to that reported for other central synapses^{17–20}, but much lower than that for refilling entire vesicle pools at a single bouton (that is, 40 s)¹⁶. Furthermore, we found that the replenishment kinetics of this pool are highly dependent on the prior pattern of activity of presynaptic terminals. High-frequency inputs

(300 Hz) induce a rapid phase of refilling that is absent when the terminal is stimulated at lower frequencies. More important, we have demonstrated that an increase in Ca^{2+} influx during repetitive firing of action potentials is the key element that enhances the replenishment. Previous studies have shown that replenishment of the readily releasable pool of granules in chromaffin cells is also a Ca^{2+} -dependent process^{25,26}. However, both Ca^{2+} influx and Ca^{2+} released from internal stores are effective in regulating replenishment. In the calyx synapse, Ca^{2+} influx through action potentials appears to be the dominant factor. Such a difference may be explained by the lack of a clearly defined active zone for release in chromaffin cells²⁷ and by the high level of Ca^{2+} -binding proteins in auditory synapses^{28,29}. These proteins may buffer rapid Ca^{2+} influx during low-frequency action-potential firing, but allow Ca^{2+} concentration to rise during high-frequency transmission. It is known that release from the docked vesicles of the readily releasable pool is tightly coupled to Ca^{2+} influx through voltage-gated Ca^{2+} channels in the active zone³⁰. Our results suggest that there is a more distant reserve pool that does not participate in immediate release but will detect changes in the level of Ca^{2+} entry and replenish the readily releasable pool accordingly. We conclude that both release and replenishment at this central synapse are highly dependent on the level of Ca^{2+} influx. Such influx may provide an important gain control mechanism to adjust synaptic strength dynamically in the central nervous system. □

Methods

Brainstem slices were prepared from postnatal mice (129SV/EMS, Jackson Laboratories) of age 12–15 days using a technique described previously^{4,5} and kept at room temperature (20–22 °C) for recordings. Whole-cell patch-clamp recordings were made from visually identified MNTB neurons or presynaptic calyces with an Axopatch 2D amplifier (for voltage clamp) or Axoclamp 2A amplifier (for current clamp) (Axon Instruments). The patch electrodes had a resistance of 2–4 M Ω for recording from postsynaptic neurons and 4–6 M Ω for recording from presynaptic terminals. EPSCs were evoked in an all-or-none fashion by stimulating the presynaptic axon fibre bundle with a platinum bipolar electrode placed near the midline of the slices. Non-somatic minor inputs can be occasionally co-activated with the main somatic calyx-type input

to the MNTB neuron (for example, small downward blips following each main EPSC in Fig. 6a). They produce small EPSCs with delayed onset, fluctuating size and failures. For recording of EPSCs, the intracellular solution contained (in mM): potassium gluconate 97.5, CsCl 32.5, EGTA 5, HEPES 10, MgCl₂ 1, ATP 2, GTP 0.2 (pH 7.2). Because of the large size of AMPA EPSCs, an intracellular Na⁺-channel blocker, QX314 (2 mM) was added to improve the quality of voltage clamp. The NMDA (N-methyl-D-aspartate) EPSC was largely blocked by Mg²⁺ (1 mM), with a residual contribution to the total EPSC at -60 mV of less than 5%. Bicuculline (10 μM) and strychnine (1 μM) were added to the artificial cerebral spinal fluid (ACSF) to block inhibitory inputs. The same intracellular solution was used for recording presynaptic action potentials and potassium currents, except that CsCl was replaced by KCl and no QX314 was added. For recordings of presynaptic calcium currents, the intracellular solution contained (in mM): CsCl 110, EGTA 0.5, HEPES 40, MgCl₂ 1, ATP 2, GTP 0.2, Na₂-phosphocreatinine 12, TEA 10 (pH 7.2) and extracellular ACSF solution was supplemented with 10 mM TEA, 0.3 mM 4-aminopyridine (4-AP) and 1 μM tetrodotoxin (TTX). The series resistance for postsynaptic recordings was 4–6 MΩ and 8–12 MΩ for presynaptic recordings. The series resistance compensation was at least 85%, with a lag of 10 μs. Data were filtered at 2–5 kHz, digitized and acquired on-line with pClamp6 software (Axon Instruments) running on a 486 computer. Current amplitude analysis and single exponential fitting were done using the same software. Averaged data are expressed as mean ± standard error (s.e.). For experiments illustrated in Fig. 6, slices were incubated in ACSF containing 0.2 mM EGTA-AM (in 0.1% dimethylsulphoxide (DMSO)) (Molecular Probes) for 30 min and then washed in normal ACSF for at least 30 min before recording. The same manipulation with DMSO alone in two control experiments produced no detectable effects. Cyclothiazide, strychnine and QX314 were purchased from Research Biochemical International; thapsigargin, 4-AP, TEA, TTX and bicuculline were from Sigma.

Received 23 February; accepted 1 May 1998.

1. Zucker, R. S. Short-term synaptic plasticity. *Annu. Rev. Neurosci.* **12**, 13–31 (1989).
2. Bliss, T. V. P. & Collingridge, G. A. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
3. Malenka, R. C. Synaptic plasticity in the hippocampus: LTP and LTD. *Cell* **78**, 535–538 (1994).
4. Forsythe, I. D. Direct patch recording from identified presynaptic terminals mediating glutamatergic EPSCs in the rat CNS, *in vitro*. *J. Physiol. (Lond.)* **479**, 381–387 (1994).
5. Takahashi, T., Forsythe, I. D., Tsujimoto, T., Barnes-Davies, M. & Onodera, K. Presynaptic calcium current modulation by a metabotropic glutamate receptor. *Science* **274**, 594–597 (1996).
6. Borst, J. G. G. & Sakmann, B. Calcium influx and transmitter release in a fast CNS synapse. *Nature* **383**, 431–434 (1996).
7. Borst, J. G. G., Helmchen, F. & Sakmann, B. Pre- and postsynaptic whole-cell recordings in the medial nucleus of the trapezoid body of the rat. *J. Physiol. (Lond.)* **489**, 825–840 (1996).
8. Kuwabara, N., DiCaprio, R. A. & Zook, J. M. Afferents to the medial nucleus of the trapezoid body and their collateral projections. *J. Comp. Neurol.* **314**, 684–706 (1991).
9. Morest, D. K. The growth of synaptic endings in the mammalian brain: a study of the calyces of the trapezoid body. *Zeitschrift für Anatomie und Entwicklungsgeschichte* **127**, 201–220 (1968).
10. Barnes-Davies, M. & Forsythe, I. D. Pre- and postsynaptic glutamate receptors at a giant excitatory synapse in rat auditory brainstem slices. *J. Physiol. (Lond.)* **488**, 387–406 (1995).
11. Trussell, L. O., Zhang, S. & Raman, I. M. Desensitization of AMPA receptors upon multiquantal neurotransmitter release. *Neuron* **10**, 1185–1196 (1993).
12. Isaacson, J. S. & Walmsley, B. Amplitude and time course of spontaneous and evoked excitatory postsynaptic currents in bushy cells of the anteroventral cochlear nucleus. *J. Neurophysiol.* **76**, 1566–1571 (1996).
13. Wang, L.-Y., Forsythe, I. D. & Kaczmarek, L. K. Regulation of high-fidelity synaptic transmission by a presynaptic potassium channel in the mouse auditory brainstem. *Soc. Neurosci. Abstr.* **23**, 365 (147.18) (1997).
14. Betz, W. J. Suppression of transmitter release at the neuromuscular junction of the frog. *J. Physiol. (Lond.)* **206**, 629–644 (1970).
15. Kusano, K. & Landau, E. M. Depression and recovery of transmission at the squid giant synapse. *J. Physiol. (Lond.)* **245**, 13–32 (1975).
16. Liu, G. & Tsien, R. W. Properties of synaptic transmission at single synaptic boutons. *Nature* **375**, 404–408 (1995).
17. Stevens, C. F. & Tsujimoto, T. Estimates for the pool size of releasable quanta at a single central synapse and for the time required to refill the pool. *Proc. Natl Acad. Sci. USA* **92**, 846–849 (1995).
18. Dobrunz, L. E. & Stevens, C. F. Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron* **18**, 995–1008 (1997).
19. von Gersdorff, H. & Matthews, G. Depletion and replenishment of vesicle pools at a ribbon-type synaptic terminal. *J. Neurosci.* **17**, 1919–1927 (1997).
20. von Gersdorff, H., Schneggenburger, R., Weis, S. & Neher, E. Presynaptic depression at a calyx synapse: the small contribution of metabotropic glutamate receptors. *J. Neurosci.* **17**, 8137–8146 (1997).
21. Guinan, J. J., Norris, B. & Guinan, S. S. Single auditory units in the superior olivary complex II: locations of unit categories and tonotopic organization. *Int. J. Neurosci.* **4**, 147–166 (1972).
22. Brownell, W. E. Organization of the cat trapezoid body and the discharge characteristics of its fibres. *Brain Res.* **94**, 413–433 (1975).
23. Wu, S. H. & Kelly, J. B. Response of neurons in the lateral superior olive and medial nucleus of the trapezoid body to repetitive stimulation: intracellular and extracellular recordings from mouse brain slice. *Hearing Res.* **68**, 189–201 (1993).
24. Parker, D. Depression of synaptic connections between identified motor neurons in the locust. *J. Neurophysiol.* **74**, 529–538 (1995).

25. Heinemann, C., Von Rüden, L., Chow, R. H. & Neher, E. A two-step model of secretion control in neuroendocrine cells. *Pflügers Arch.* **424**, 105–112 (1993).
26. Von Rüden, L. & Neher, E. A Ca-dependent early step in the release of catecholamines from adrenal chromaffin cells. *Science* **262**, 1061–1065 (1993).
27. Hankel, A. W. & Almers, W. Fast steps in exocytosis and endocytosis studied by capacitance measurements in endocrine cells. *Curr. Opin. Neurobiol.* **6**, 350–357 (1996).
28. Vater, M. & Braun, K. Parvalbumin, calbindin D-28k, and calretinin immunoreactivity in the ascending auditory pathway of horseshoe bats. *J. Comp. Neurol.* **341**, 534–558 (1994).
29. Lohmann, C. & Friauf, E. Distribution of the calcium-binding proteins parvalbumin and calretinin in the auditory brainstem of adult and developing rats. *J. Comp. Neurol.* **367**, 90–109 (1996).
30. Llinas, R., Sugimori, M. & Silver, R. B. Microdomains of high calcium concentration in a presynaptic terminal. *Science* **256**, 677–679 (1992).

Acknowledgements. We thank I. Forsythe for his helpful insight and for sharing unpublished data, and N. Magoski, M. Whim and W. Joiner for critically reading the manuscript. This work is supported by an NIH grant to L.K.K. and a postdoctoral fellowship from the Epilepsy Foundation for Research to L.-Y.W.

Correspondence and requests for materials should be addressed to L.K.K. (e-mail: Kaczmarek@Yale.edu).

An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy

Kazuhiro Kobayashi^{††}, Yutaka Nakahori[‡], Masashi Miyake[†], Kiichiro Matsumura[§], Eri Kondo-Iida^{*||}, Yoshiko Nomura[¶], Masaya Segawa[¶], Mieko Yoshioka[#], Kayoko Saito^{||}, Makiko Osawa^{||}, Kenzo Hamano[☆], Youichi Sakakihara^{**}, Ikuya Nonaka^{††}, Yasuo Nakagome[†], Ichiro Kanazawa^{‡‡}, Yusuke Nakamura^{§§}, Katsushi Tokunaga[†] & Tatsushi Toda^{*}

Laboratories of ^{*}Genome Medicine, and ^{§§}Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

Departments of [†]Human Genetics, ^{**}Pediatrics, and ^{‡‡}Neurology, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan

[‡] Department of Public Health, School of Medicine, University of Tokushima, Tokushima 770-8503, Japan

[§] Department of Neurology, Teikyo University School of Medicine, Tokyo 173-8605, Japan

^{||} Department of Pediatrics, Tokyo Women's Medical College, Tokyo 162-8666, Japan

[¶] Segawa Neurological Clinic for Children, Tokyo 101-0062, Japan

[☆] Department of Pediatrics, Kobe General Hospital, Kobe 650-0046, Japan

[#] Department of Pediatrics, Institute of Clinical Medicine, University of Tsukuba, Tsukuba 305-8576, Japan

^{††} Department of Ultrastructural Research, National Institute of Neuroscience, NCNP, Tokyo 187-8502, Japan

Fukuyama-type congenital muscular dystrophy (FCMD), one of the most common autosomal recessive disorders in Japan (incidence is 0.7–1.2 per 10,000 births), is characterized by congenital muscular dystrophy associated with brain malformation (micro-polygria) due to a defect in the migration of neurons¹. We previously mapped the FCMD gene to a region of less than 100 kilobases which included the marker locus *D9S2107* on chromosome 9q31 (refs 2–4). We have also described a haplotype that is shared by more than 80% of FCMD chromosomes, indicating that most chromosomes bearing the FCMD mutation could be derived from a single ancestor⁵. Here we report that there is a retrotransposal insertion of tandemly repeated sequences within this candidate-gene interval in all FCMD chromosomes carrying the founder haplotype (87%). The inserted sequence is about 3 kilobases long and is located in the 3' untranslated region of a gene encoding a new 461-amino-acid protein. This gene is expressed in various tissues in normal individuals, but not in FCMD patients who carry the insertion. Two independent point mutations confirm that mutation of this gene is responsible for FCMD. The predicted protein, which we term fukutin, contains an amino-terminal signal sequence, which together with results from

transfection experiments suggests that fukutin is a secreted protein. To our knowledge, FCMD is the first human disease to be caused by an ancient retrotransposal integration.

After localizing the FCMD gene to the centromeric vicinity of the *D9S2170* locus by linkage-disequilibrium mapping⁴ and founder haplotype mapping⁵, we constructed a cosmid contig harbouring *D9S2105* and *D9S2107* (ref. 6). To screen DNA from FCMD patients for genomic rearrangement, we hybridized each cosmid clone of the contig (Fig. 1a) to Southern blots of genomic DNA from FCMD patients. When the whole insert of cosmid clone cE6, which included *D9S2170*, was used as a probe, we found that FCMD DNAs digested with *PvuII* specifically lacked the normal 5-kilobase (kb) band and instead had a new 8-kb band. When the same DNAs were digested with *PstI* or *BglII*, the normal band was missing and a new, larger band appeared, indicating that a 3-kb fragment had been inserted into the genomic DNA of most of our FCMD patients (Fig. 1b). When a fragment of this cosmid, E6f3, was used as a probe, most patients showed an abnormal 8-kb homozygous band; the remaining patients gave both a normal 5-kb and an abnormal 8-kb

band (Fig. 1c). Detailed analysis confirmed the insertion of a 3-kb sequence within a 1.4-kb *EcoRI* fragment of cE6 in most FCMD patients (Fig. 1a). This insertion allele cosegregated with the FCMD founder haplotype (138-192-147-183 for *D9S2105-D9S2170-D9S2171-D9S2107*).

Using the 1.4-kb *EcoRI* fragment of cE6 as a probe to screen an adult brain complementary DNA library, we obtained 2.1-kb and 3.2-kb cDNAs with poly(A) tracts. Screening of additional cDNA libraries or RACE (rapid amplification of cDNA ends) experiments yielded 13 additional clones. Sequencing of all 15 cDNA clones revealed that the composite cDNA spans 7,349 base pairs (bp) and an open reading frame of 1,383 bp begins with an initiator ATG codon at base 112. The predicted protein has 461 amino acids, a calculated relative molecular mass of 53.6K and a calculated isoelectric point of 8.29 (Fig. 2). The original two cDNA clones resulted from polyadenylation at two consensus sites (bases 6,266 and 7,327), and corresponded to the two bands seen on northern blot (Fig. 3).

Northern blots revealed 6.5-kb and 7.5-kb transcripts in poly(A)⁺ messenger RNA from a variety of human tissues, with the strongest signals being obtained from heart, brain, skeletal muscle and pancreas. This may explain why congenital muscular dystrophy in FCMD patients is associated with a brain anomaly and often with fibrosis of the heart muscle as well¹. Transcripts were also detected in cultured lymphoblasts (Fig. 3).

The genomic structure of this gene was determined by Southern hybridization against cosmid clones using as probes DNA sequences corresponding to each exon, and by DNA sequencing. The gene spans more than ~100 kb of genomic DNA and is composed of 10 exons; exon 2 includes the translation-initiation codon (Fig. 1a).

Searches of the nucleotide databases revealed no significant similarities to genes of known function. The predicted protein sequence was compared against the SwissProt and PIR databases, and against the predicted translation products of GenBank, EMBL, DDBJ and PDB databases. No significant similarity was found to any known proteins, although weak similarities to the predicted translation products from *Caenorhabditis elegans* cosmids W02B3 and T07A4 (*U22833* and *Z48055*) were detected. A search for protein motifs revealed one N-glycosylation site. The PSORT program⁷ for prediction of protein localization sites indicated that the FCMD protein has an N-terminal signal sequence with a possible cleavage site at codon 21, but that it lacks a transmembrane region; another program, TMpred⁸, however, predicted one (amino acids 288–306) transmembrane segment (Fig. 2).

To characterize the inserted DNA, we isolated a clone containing the insertion from a genomic library constructed from an FCMD patient who was homozygous for the insertion. Sequence analysis revealed that the inserted DNA fragment, which was 3,062 bp long, was composed of (TCTCCC)₄₁; 27 copies of a 49-bp sequence (5'-GGGAGGGAGGTGGGGGGTACGCCCCCG-CCTGGCCAGCCGCCCATCC-3', tandemly repeated); a SINE (short interspersed sequence)-type human retroposon sequence;

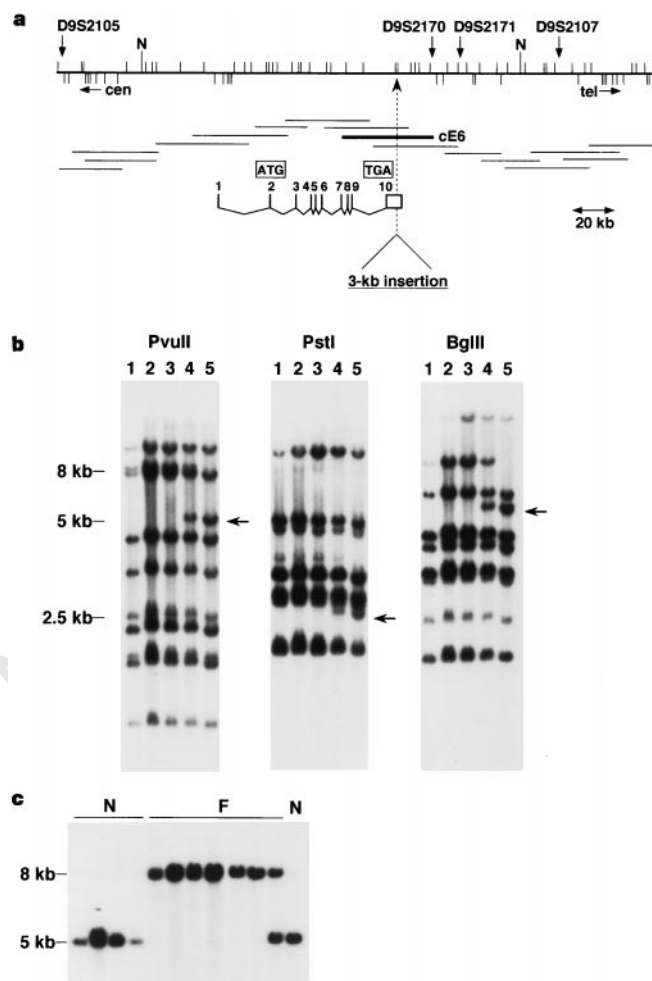


Figure 1 Genomic insertion in FCMD patients. **a**, The FCMD gene region on chromosome 9q31. Horizontal lines, individual cosmids of the cosmid contig; short vertical lines, *EcoRI* cutting sites; N, *NotI* cutting sites. The genomic structure of the FCMD gene is shown below the map. **b**, Southern blot analysis of FCMD patients. Genomic DNA was digested with *PvuII*, *PstI* or *BglII*. Lanes: 1–3, patients homozygous for the founder haplotype; 4, patient heterozygous for the founder haplotype; 5, a normal control. The probe used was the whole insert of cosmid clone cE6. Bands seen in the normal control were absent in FCMD patients (arrows) and new bands larger than the missing band were evident. **c**, *PvuII*-digested blot hybridized with a fragment probe, E6f3, derived from cE6. F, FCMD patients; N, normal controls.

1 MSRKNNVFLALLTLTSSAFLLFQLYYYKHLYSTKNGAGLSKSGSRIGF
51 DSTQWRAVKKFIIMLTSTNQNVFVLIDPLILELINKNFEQVKNTSHGSTSQ
101 CKFFCVPRDFTAFALQYHLWKNEEGWFRIAENMGFQCLKIESKDPRLDGI
151 DLSLSTGEIPLHYICKLATHAHLVVFHERSGNYLWHGHLRLKEHIDRKFV
201 PFQKLQFGRYPGAFDRPELQVTVVDGLEVLIPKDPMHFVEEVPHSRFTEC
251 RYKEARAFFQOYLDDNTVEAVAFRKSARELLQAAKTLNKLGVFPFLSSG
301 TCLGWYRQCNIIPYSKDVLDLGIQDYKSDIILAFQDAGLPLKHKFGKVE
351 DSLELSFQKDDVKLDVFFVEETDHMMWNGGTQAKTGKFKYLPFKFTLC
401 WTEFVDMKVHVPCETLEYIEANYGKTKWIPVKTWDRKSPNVQPNGIWP
451 ISEWDEVIQLY

Figure 2 Amino-acid sequence of the protein (fukutin) encoded by the FCMD gene. The predicted hydrophobic signal sequence is underlined. A potential N-linked glycosylation site is indicated by a broken line. Amino acids mutated in the families described here are indicated by a hash symbol.

a polyadenylation signal (AATAAA); and poly(A). Furthermore, a target-site duplication, consisting of a direct repeat of AAGAAAAAATTGT at both ends, indicated retrotransposal insertion of this 3-kb fragment. The tandem-repeat sequence was almost identical to one observed in the human gene encoding complement component 2 (Z11739), the genomic region near the Huntington's disease gene (Z69654).

By comparing cDNA and genomic sequences, we conclude that the 3-kb insertion was in the 3' untranslated region of this gene (between bases 5,889–5,890). Most (125 of 144, 87%) FCMD chromosomes contained this insertion, whereas it was found in only one of 176 chromosomes in unrelated normal individuals; the frequency of one in 88 individuals corresponded well to that of FCMD carriers in the Japanese population¹.

To determine the effect of this genomic insertion on the FCMD transcript, we did northern blot analysis of poly(A)⁺ mRNA isolated from cultured lymphoblasts of several patients. The transcript of this gene was nearly undetectable in FCMD patients who carried the insertion homozygously, and significantly lower than normal in patients heterozygous for the insertion and another mutation haplotype (Fig. 3b). The results indicated that the insertion or other mutations in the FCMD gene are likely to cause a decrease in transcription and/or an instability of mRNA, resulting in loss of function.

To search for inactivating mechanisms in FCMD chromosomes without the 3-kb insertion, we screened exons and flanking intronic sequences throughout the coding region by using single-stranded conformation polymorphism (SSCP) analysis in four patients. A product of the polymerase chain reaction (PCR) corresponding to exon 3 of the gene in patient H.M. (Fig. 4a, left), who carried the founder insertion from her father and the 130-201-157-183 haplotype from her mother (corresponding to lane 3 of the northern blot shown in Fig. 3b), showed a mobility shift. Sequence analysis showed a C-to-T transition at base 250 in the maternal allele, resulting in premature termination (CGA to TGA, R47X; Fig. 4b, left). Analysis of other members of H.M.'s family by PCR and restriction-fragment length polymorphism (RFLP) indicated that this mutation segregated with the disease and with the microsatellite haplotype. Furthermore, five other families (those of Y.S., R.M., A.G., A.T. and A.Y.) carrying the same haplotype had the same

nonsense mutation (Fig. 4c, left). In addition, we identified in patient T.I. a 2-bp deletion at bases 298–299 (codon 63) causing a frameshift and a premature stop at codon 75 (Fig. 4a, b, right). The patient's Japanese mother carries an insertion-chromosome; her American father (of English and German extraction) does not. The 2-bp deletion was present in the chromosome inherited from her father (Fig. 4c, right). Taking together, our results indicate that our cDNA and its predicted protein, fukutin, are the FCMD gene and its product.

To investigate the localization of fukutin in skeletal muscle, we raised six polyclonal and three monoclonal antibodies; however, none of these detected specific signals on either control or FCMD skeletal muscle (data not shown). Computer analysis suggested that fukutin could be a membrane protein or an extracellular protein. Next, we transfected COS-7 cells with plasmids carrying fukutin-GFP (where GFP is green fluorescent protein) or fukutin-HA (HA is haemagglutinin). Transfected cells gave an intracellular signal that revealed a perinuclear ribbon-like Golgi apparatus (Fig. 5a); the labelling patterns of each transfectant were almost identical. Fukutin co-localized with a Golgi marker, Golgi 58K protein (Fig. 5b, c). Gradually, the signals assume a granular cytoplasmic distribution (Fig. 5d), suggesting that fukutin passes through the Golgi before being packaged into secretory vesicles. The signal was not seen at the plasma membrane, however, where most proteins responsible for the muscular dystrophies are located. Transfection experiments using mouse myoblast cell line C2C12 cells gave similar results.

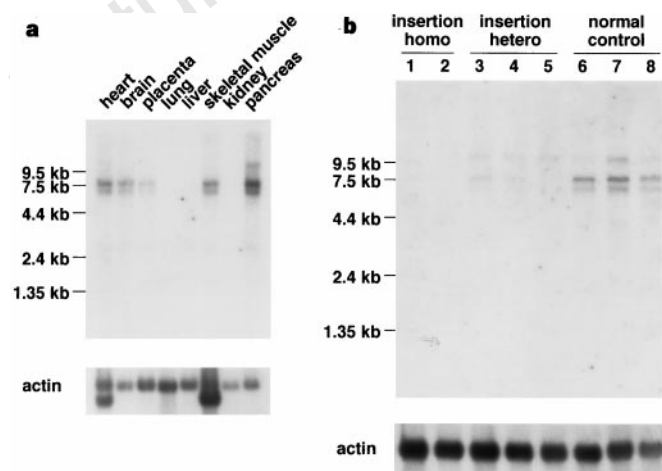


Figure 3 Northern blot analysis of the FCMD gene. **a**, Tissue-specific expression of FCMD mRNA in human tissues. An RNA blot containing 2 μ g poly(A)⁺ RNA from each of eight human tissues was hybridized with FCMD cDNA (clone II-5 containing the coding sequence) and β -actin (shown below) cDNA probes. **b**, RNA blot containing 2 μ g poly(A)⁺ RNA from lymphoblasts of FCMD patients and normal controls, hybridized with the same probes as in **a**. Note that the signals are almost absent or abnormally low in FCMD patients who carry the insertion homozygously or heterozygously, respectively.

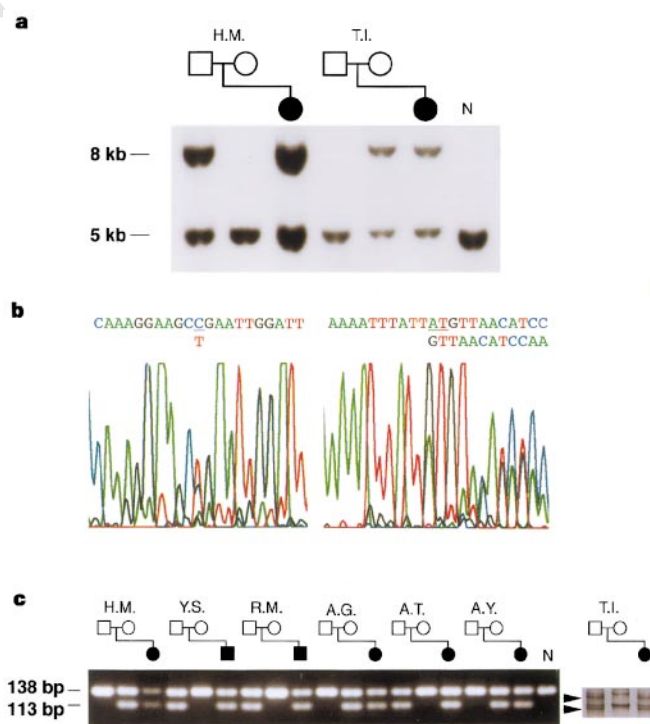


Figure 4 Segregation of point mutations in the families of H.M. and T.I. **a**, Patients H.M. and T.I. are heterozygous for the founder-insertion allele. Patient T.I. is the daughter of an American father and Japanese mother. **b**, Patients were shown to carry C250T (left) and Δ 298–299 (right) mutations, both of which cause premature termination (R47X and frameshift). **c**, *AluI*, which cleaves DNA at AGCT sequences, was used to confirm the mutation in the family of H.M. and in five other families (those of Y.S., R.M., A.G., A.T. and A.Y.) carrying the same haplotype (left); the mutant allele has an AGCT sequence. In this case, the *AluI* digests of the normal PCR product are 138 bp and 4 bp; *AluI* digestion of the mutant allele divided it into 113-bp, 25-bp and 4-bp fragments. In the family of H.M. the (non-insertion) mutant allele is derived from the mother. SSCP analysis showed that the 2-bp deletion in patient T.I. was derived from her caucasian father (right, arrowheads).

To confirm that fukutin is a secreted protein, we looked for it in the culture medium of COS-7 cells expressing fukutin-GFP by using metabolic labelling and immunoprecipitation with an anti-GFP polyclonal antibody. This antibody specifically precipitated fukutin from medium from COS-7 cells transfected with fukutin-GFP but not from control culture medium (Fig. 5e), indicating that fukutin-GFP was being secreted into the COS-7 cell culture medium. We conclude that fukutin is a secreted protein rather than a plasma-membrane protein.

Our results are evidence that the principal FCMD ancestor chromosome carries a 3-kb insertion within the 3' untranslated region that results in dysfunction of the FCMD gene. How can insertion into the 3' untranslated region cause a reduction in the corresponding mRNA? As the 3' untranslated region affects the stability of mRNA⁹, the inserted sequence may alter its secondary structure and render it unstable. Defects in the dystrophin-glycoprotein complex¹⁰ lead to a loss of linkage between laminin $\alpha 2$ in the extracellular matrix and actin in the subsarcolemmal cytoskeleton, cause various muscular dystrophies^{11,12}. Decreased immunostaining

of a sarcolemmal glycoprotein, β -dystroglycan¹³, and of an extracellular matrix protein linked with α -dystroglycan, the laminin $\alpha 2$ chain¹⁴, has been observed in FCMD muscle. Abnormalities in basal lamina in FCMD muscle and brain have been seen by electron microscopy^{15,16}. In light of these observations, we suggest that fukutin may be located in the extracellular matrix, where it interacts with and reinforces a large complex encompassing the outside and inside of muscle membranes; alternatively, as a secreted protein, fukutin may cause muscular dystrophy by an unknown mechanism. Another important manifestation of FCMD is a brain anomaly called micropolygyria (type II lissencephaly), in which neuronal lamination of normal six-layered cortex is lacking because of a defect in the migration of neurons¹. At least four genes implicated in cortical dysgenesis disorders have been proposed to function in the migration and assembly of neurons during cortical histogenesis¹⁷⁻²¹. Of these, like fukutin, reelin is a secreted glycoprotein that has several structural features characteristic of extracellular matrix proteins. Discovery of the FCMD gene represents an important step towards greater understanding of the pathogenesis of muscular dystrophies and also of normal brain development. □

Methods

FCMD families and Southern/northern blot analysis. DNA samples from 63 families²⁻⁴ and 23 additional FCMD families were analysed. Southern blots were analysed by competitive hybridization. Multiple-tissue northern blots were purchased from Clontech. Total RNA was isolated from lymphoblasts of patients and normal controls by extraction with acid-phenol (Trizol, Gibco-BRL). Poly(A)⁺ RNA (2 μ g) was selected by oligo(dT) chromatography (Oligo-dT Latex, Takara).

Screening of cDNA libraries, RACE and sequence analysis. A cDNA library from human adult brain (caudate nucleus) was supplied by Y. Misumi. Libraries from fetal brain were purchased from Clontech. RACE experiments were performed using the Marathon cDNA amplification kit and Marathon Ready cDNA from adult heart (Clontech) in parallel. Sequencing was done using the PRISM ReadyReaction DyeDeoxy Terminator cycle-sequencing kit with AmpliTaq-FS (Perkin-Elmer) on an ABI Model 377 DNA sequencer (Perkin-Elmer).

SSCP analysis. Primers ex3F (5'-GTTGCATGCTGGACTTTGAA-3') and ex3R (5'-CATAAAGCACTTGGTAAAGGGC-3') for exon 3, and primers ex4F (5'-GACTGTGTGTTGGCTTACTGG-3') and ex4R1 (5'-GATACTGCAGTGCAAATGCAG-3') for exon 4, were used to amplify patient DNA. SSCP analysis was performed as described²².

Cell transfection and immunofluorescence. A cDNA fragment (bases 60-1,494) containing an open reading frame was subcloned by PCR into the vector pEGFP-N1 (Clontech) and the vector pcDNA3.1(+) (Invitrogen). On fukutin-HA, an HA epitope tag (YPYDVPDYA) was added to the C terminus of fukutin. COS-7 and C2C12 cells were transfected with plasmid DNA with LipofectAMINE (Gibco-BRL) in Opti-MEM (Gibco-BRL). The localization of fukutin-GFP, fukutin-HA and Golgi 58K protein was determined by fluorescence microscopy one day after transfection.

Metabolic labelling and immunoprecipitation. Fukutin-GFP expressed COS-7 cells were incubated in DMEM containing [³⁵S]methionine and [³⁵S]cysteine. Culture medium was collected and pre-cleared by centrifugation. Anti-GFP polyclonal antibody (Clontech) was added to the samples. Immunoprecipitates were captured with protein A-Sepharose-4b (Zymed) and washed with RIPA buffer (10 mM Tris-HCl, pH 7.4, 1% Nonidet-P40, 0.1% Na deoxycholate, 0.1% SDS and 150 mM NaCl). Samples were electrophoresed on 10% SDS-polyacrylamide gels.

Received 20 April; accepted 1 June 1998.

1. Fukuyama, Y., Osawa, M. & Suzuki, H. Congenital muscular dystrophy of the Fukuyama type—clinical, genetic and pathological considerations. *Brain Dev.* **3**, 1–30 (1981).
2. Toda, T. *et al.* Localization of a gene for Fukuyama type congenital muscular dystrophy to chromosome 9q31–33. *Nature Genet.* **5**, 283–286 (1993).
3. Toda, T. *et al.* Refined mapping of a gene responsible for Fukuyama-type congenital muscular dystrophy: evidence for strong linkage disequilibrium. *Am. J. Hum. Genet.* **55**, 946–950 (1994).
4. Toda, T. *et al.* Linkage-disequilibrium mapping narrows the Fukuyama-type congenital muscular dystrophy (FCMD) candidate region to <100 kb. *Am. J. Hum. Genet.* **59**, 1313–1320 (1996).
5. Kobayashi, K. *et al.* Founder-haplotype analysis in Fukuyama-type congenital muscular dystrophy (FCMD). *Hum. Genet.* (in the press).

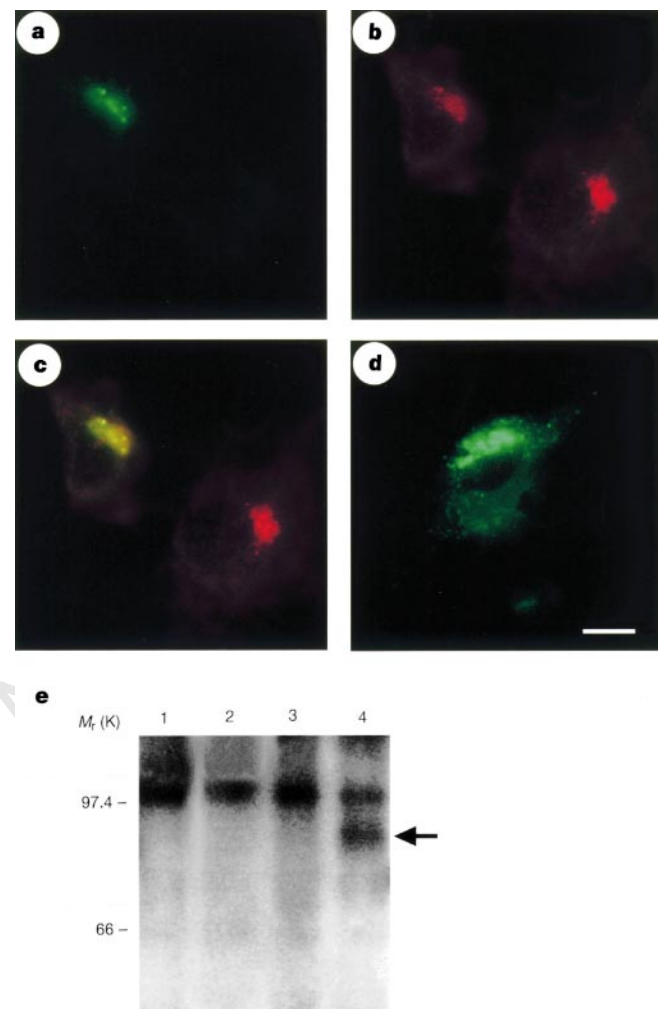


Figure 5 Cellular localization of fukutin. **a**, COS-7 cells transfected with a fukutin-GFP construct show a perinuclear ribbon-like signal (green). **b**, Immunolocalization of the same cell with anti-Golgi 58K protein monoclonal antibody (red) reveals an almost identical localization with fukutin-GFP. **c**, Superimposition of the two labelling methods. **d**, C2C12 cells transfected with a fukutin-GFP construct; the signal appears to have a granular cytoplasmic distribution. Scale bar, 10 μ m. **e**, Secretion of fukutin. COS-7 cells were transfected with fukutin-GFP (lanes 3 and 4) or without it (lanes 1 and 2). The ³⁵S-labelled supernatant was immunoprecipitated with no antibody (lanes 1 and 3) or with anti-GFP antibody (lanes 2 and 4). A specific band of M_r ~86K (arrow) was detected only in the sample containing both fukutin-GFP and anti-GFP.

- Miyake, M. *et al.* YAC and cosmid contigs encompassing the Fukuyama-type congenital muscular dystrophy (FCMD) candidate region on 9q31. *Genomics* **40**, 284–293 (1997).
- Nakai, K. & Kanehisa, M. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**, 897–911 (1992).
- Hofmann, K. & Stoffel, W. TMbase—A database of membrane spanning protein segments. *Biol. Chem. Hoppe Seyler* **374**, 166 (1993).
- Sachs, A. B. Messenger RNA degradation in eukaryotes. *Cell* **74**, 413–421 (1993).
- Ervasti, J. M. & Campbell, K. P. Membrane organization of the dystrophin–glycoprotein complex. *Cell* **66**, 1121–1131 (1991).
- Campbell, K. P. Three muscular dystrophies: loss of cytoskeleton–extracellular matrix linkage. *Cell* **80**, 675–679 (1995).
- Worton, R. Muscular dystrophies: disease of the dystrophin–glycoprotein complex. *Science* **270**, 755–756 (1995).
- Matsumura, K., Nonaka, I. & Campbell, K. P. Abnormal expression of dystrophin-associated proteins in Fukuyama-type congenital muscular dystrophy. *Lancet* **341**, 521–522 (1993).
- Hayashi, K. Y. *et al.* Abnormal localization of laminin subunits in muscular dystrophies. *J. Neurol. Sci.* **119**, 53–64 (1993).
- Nakano, I., Funahashi, M., Takada, K. & Toda, T. Are beaches in the glia limitans the primary cause of the micropolygia in Fukuyama-type congenital muscular dystrophy (FCMD)?—Pathological study of the cerebral cortex of an FCMD fetus. *Acta Neuropathol.* **91**, 313–321 (1996).
- Ishii, H., Hayashi, Y. K., Nonaka, I. & Arahata, K. Electron microscopic examination of basal lamina in Fukuyama congenital muscular dystrophy. *Neuromuscul. Disord.* **7**, 191–197 (1997).
- D'Arcangelo, G. *et al.* A protein related to extracellular matrix proteins deleted in the mouse mutant reeler. *Nature* **374**, 719–723 (1995).
- Sheldom, M. *et al.* Scrambler and *yotari* disrupt the *disabled* gene and produce a *reeler*-like phenotype in mice. *Nature* **389**, 730–733 (1997).
- Reiner, O. *et al.* Isolation of a Miller-Dieker lissencephaly gene containing G-protein β -subunit-like repeats. *Nature* **364**, 717–721 (1993).
- des Portes, V. *et al.* A novel CNS gene required for neuronal migration and involved in X-linked subcortical laminar heterotopia and lissencephaly syndrome. *Cell* **92**, 51–61 (1998).
- Gleeson, J. G. *et al.* *doublecortin*, a brain-specific gene mutated in human X-linked lissencephaly and double cortex syndrome, encodes a putative signaling protein. *Cell* **92**, 63–72 (1998).
- Bannai, M. *et al.* Discrimination of human HLA-DRB1 alleles by PCR-SSCP (single-strand conformation polymorphism) method. *Eur. J. Immunogenet.* **21**, 1–9 (1994).

Acknowledgements. We thank the family members who participated in this study; F. Saito, Y. Sunada, E. Katayama, M. Hino, J. Fujisawa, F. Oyama, I. Nishino, H. Hohjoh, Y. Miki, M. Watanabe, T. Tokino, K. Arahata, Y. Fukuda, J. Inazawa, T. Tanaka and S. Ikegawa for advice; T. Iwata, T. Kumagai, S. Kobayashi, T. Ishihara, K. Suzumori, C. Nakano, S. Sugino, M. Funahashi, Y. Mizuno, H. Muranaka, A. Honma and E. Nanba for samples; J. Goto and N. Tsuji for managing cell lines; and I. Matsushita and K. Takeuchi for technical assistance. This work was supported by grants from the Ministry of Health and Welfare and from the Ministry of Education, Science and Culture, Japan.

Correspondence and requests for materials should be addressed to T.T. (e-mail: toda@ims.u-tokyo.ac.jp). The GenBank/EMBL/DBJ accession number for the sequence reported in this Letter is AB008226.

Developmental selection of *var* gene expression in *Plasmodium falciparum*

Qijun Chen^{*†‡}, Victor Fernandez^{*‡}, Annika Sundström^{*}, Martha Schlichtherle^{*}, Santanu Datta[§], Per Hagblom^{*} & Mats Wahlgren^{*}

^{*} Microbiology and Tumor Biology Center, Karolinska Institutet and the Swedish Institute for Infectious Disease Control, Box 280, S-171 77 Stockholm, Sweden

[†] Changchun University of Agriculture and Animal Sciences, 175 Xi An Da Lu, 130062 Changchun, People's Republic of China

[§] Astra Research Centre, P.B. 359, Mälleswaram, Bangalore, India

[‡] These authors contributed equally to this work.

The protozoan *Plasmodium falciparum* causes lethal malaria¹. Adhesion of erythrocytes infected with *P. falciparum* to vascular endothelium and to uninfected red blood cells (rosetting) may be involved in the pathogenesis of severe malaria^{2–4}. The binding is mediated by the antigenically variant erythrocyte-membrane-protein-1 (PfEMP-1)^{5–8}, which is encoded by members of the *P. falciparum var* gene family^{9,10}. The control of expression and switching of *var* genes seems to lack resemblance to mechanisms operating in variant gene families of other microbial pathogens^{11,12}. Here we show that multiple, distinct *var* gene transcripts (about 24 or more) can be detected by reverse transcription and polymerase chain reaction in bulk cultures of the rosetting parasite FCR3S1.2, despite the adhesive homogeneity of the cultures. We also detected several *var* transcripts in single erythrocytes infected with a ring-stage parasite of FCR3S1.2, and found that different *var* genes are transcribed simultaneously

from several chromosomes in the same cell. In contrast, we detected only one *var* transcript, FCR3S1.2 *var*-1, which encodes the rosetting PfEMP-1 protein¹³, in individual rosette-adhesive trophozoite-infected cells, and we found only one PfEMP-1 type at the erythrocyte surface by labelling with ¹²⁵I-iodine and immunoprecipitation. We conclude that a single *P. falciparum* parasite simultaneously transcribes multiple *var* genes but, through a developmentally regulated process, selects only one PfEMP-1 to reach the surface of the host cell.

We studied the transcription and expression of *var* genes in a recently cloned and phenotypically homogeneous parasite culture, FCR3S1.2, which has a rosetting rate of greater than 90%. We first used reverse transcription and polymerase chain reaction (RT-PCR) to amplify the *var* gene messenger RNAs from unsynchronized cultures at the eighteenth generation after cloning, when most of the parasites were at the ring stage. We used degenerate primers that mapped to conserved stretches of the Duffy-binding-like (DBL)-1 domain of PfEMP-1 (ref. 14) (Fig. 1). We cloned products of amplification by RT-PCR, and sequenced 15 of these recombinant plasmids containing inserts of 500–600 base pairs (bp). As the parasites were phenotypically homogeneous, we assumed that there would be a corresponding sequence homogeneity in the transcript population. Analysis of the inserts showed, however, that only ~30% of the sequences (5 of 15) corresponded to the FCR3S1.2

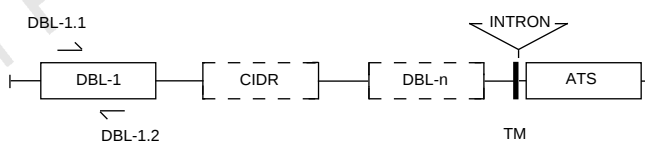


Figure 1 General structure of *var*/PfEMP-1. The semi-conserved Duffy-binding-like (DBL)-1 domain is present in all *var* genes sequenced to date. Downstream of the DBL-1 domain there is a cysteine-rich interdomain region (CIDR) and a variable number ($n = 0-4$) of less conserved DBL structures. A putative transmembrane domain (TM) is followed by the highly conserved acidic terminal sequence (ATS), which is presumably cytoplasmically located. We designed the degenerate oligonucleotide primers DBL-1.1 and DBL-1.2 from short, conserved amino-acid sequences flanking relatively variable stretches (400–700 bp) in the DBL-1 domain¹⁴. Boxes with dashed outlines indicate that these domains may or may not be present in a given PfEMP-1 molecule.



Figure 2 Predicted amino-acid sequences of *var* transcripts from bulk cultured FCR3S1.2. The figure shows the alignment of the unique sequences (*var* 1–8) that were amplified by RT-PCR with the DBL-1.1 and DBL-1.2 primers. Amino-acid similarity is indicated by boxes. Gaps and incomplete sequences are represented by dashed and dotted lines, respectively.

var-1 PfEMP-1, which is the rosetting ligand of this parasite¹³. Ten out of 15 recombinants, of which 8 were unique, contained *var* sequences that were distinct from FCR3S1.2 *var-1* (Fig. 2). All of the amplified products encoded DBL-1 domains. In another set of RT-PCR amplifications of total RNA, a similar pattern of transcript diversity was observed in the rosetting-negative FCR3S/a, a related but phenotypically distinct parasite (data not shown). Thus we found a high degree of sequence diversity of active *var* genes (and presumably, therefore, functional variation) in these bulk cultures.

Were the multiple transcripts a consequence of the extensive antigenic switching of the *var* genes, or do individual parasites express several PfEMP-1 mRNAs? To answer this question we amplified the PfEMP-1 mRNAs from individual ring-infected cells, trophozoite-infected cells, and uninfected erythrocytes from the same FCR3S1.2 culture by using single-cell micromanipulation cloning and subsequent RT-PCR. The conditions for the RT-PCR with the DBL-1 primers were optimized for use with a minimal amount of template¹³. The results show that more than one amplified product was present in at least seven out of fourteen

ring-stage parasites, as seen by gel analysis (Fig. 3a). The most frequent RT-PCR product was a band of ~430 bp. This amplification product (see below) was also detected in the bulk culture when using the optimized RT-PCR conditions. In addition to the 430-bp band, a 300-bp product was seen in six of the fourteen cells; a 550-bp amplification product was found in five cases and a 750-bp band was detected in one of the *P. falciparum*-infected red blood cells. Importantly, RT-PCR amplification products were not observed in reactions from which reverse transcriptase was omitted, nor was intron DNA amplified when using primers spanning the intron/exon splice junction of FCR3S1.2 *var-1* (Fig. 3b, c).

We cut out of the gel PCR products from several cells (430- and 550-bp bands; Fig. 3a). We then reamplified, cloned and sequenced these products. The 550-bp bands showed sequence heterogeneity. Sequencing of four individual 550-bp clones from cell 3 showed a new variant of the DBL-1 domain (*var-9*) and three sequences identical to FCR3S1.2 *var-1* (Fig. 3d). We found two additional DBL-1 variants (*var-10* and *var-11*) when we sequenced two 550-bp clones from cell 7. All of these DBL-1 variants identified in individual cells, except *var-11*, were found in the FCR3S1.2 bulk culture. As in the bulk culture, FCR3S1.2 *var-1* DBL-1 was the most common transcript sequence in single parasites. The data show that several *var* genes are transcribed in individual *P. falciparum* parasites early during intra-erythrocytic development.

In contrast to the sequence diversity found in the 550-bp bands, sequencing of the 430-bp RT-PCR products from five ring-infected red blood cells (cells 1, 2, 5, 7 and 9) showed that this band contained only one *var* sequence (of 434 bp) that was identical to the sequence of the original FCR3S1.2 *var-1* gene of this parasite. We have shown that the RT-PCR products from rosetting, trophozoite-infected red blood cells containing FCR3S1.2, always consist of the same 434-bp band containing only the *var-1* sequence¹³. We have confirmed this finding independently (not shown) and it concurs with the detection of a single *var* transcript in mature parasites of a homogeneous phenotype¹⁵. The results indicate that early expression of multiple *var* genes is curtailed to a monogenic expression of PfEMP-1 as the parasite progresses from the ring to the trophozoite stage, between 8 and 16 hours after invasion. Although the mechanism of selection remains obscure, the fact that *var-1* is the most frequently expressed transcript in individual parasites, despite being in a minority in the total *var* message population at the ring stage, suggests that the selection process occurs post-transcriptionally. Other layers of control, for example at the transcriptional level, may also be in operation.

To further study the transcription of *var* genes in single parasites,

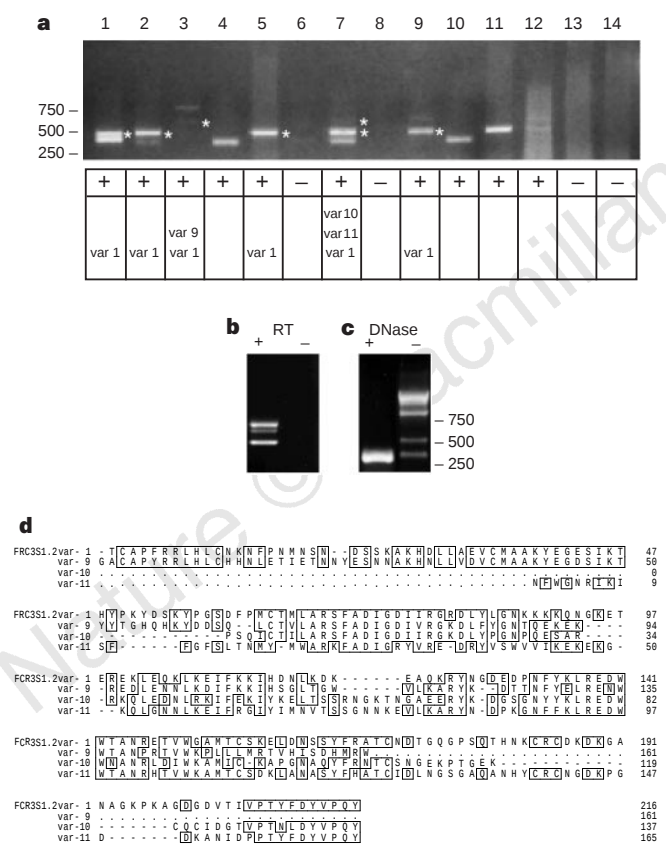


Figure 3 Identification of *var* transcripts in single *P. falciparum* parasites at the ring stage. **a**, Products amplified by RT-PCR with primers DBL-1.1 and DBL-1.2 from individual ring-infected red blood cells (cells 1–14) were separated by agarose gel electrophoresis. Bands marked with asterisks were cut out from the gel, reamplified with the same primers, cloned and sequenced. The presence (+) or absence (–) of detectable amplified products, as well as the number and type of *var* transcripts, is given in boxes below each individual cell lane. Molecular sizes are indicated in base pairs. **b**, RT-PCR amplification with primers DBL-1.1 and DBL-1.2 of RNA from *P. falciparum*-infected red blood cells (pRBCs) in the presence or absence of reverse transcriptase (RT). **c**, RT-PCR amplification with primers P-1 and L-5.1, which span the intron/exon junction site of FCR3S1.2 *var-1*. Reactions were performed with or without DNase treatment. **d**, Alignment of the predicted amino-acid sequences of *var-1* and *var-9–11* RT-PCR-amplified from single pRBCs. Amino-acid similarity is indicated by boxes. Gaps and incomplete sequencing are represented by dashed and dotted lines, respectively.

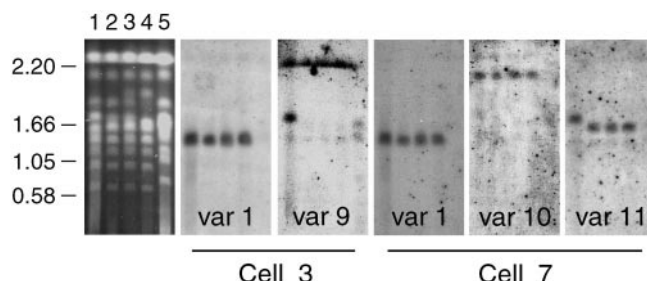


Figure 4 Chromosomal hybridization of *var* genes simultaneously expressed in individual parasites. Chromosomes of *P. falciparum* clones and lines (lane 1, FCR3S1.6 (rosetting-negative); lane 2, FCR3S1.2 (rosetting-positive); lane 3, FCR3S1 (rosetting-positive); lane 4, FCR3S/a (rosetting-negative) and lane 5, 3D7A (rosetting-negative) were separated by pulsed field gel electrophoresis, transferred to filters and hybridized with the indicated *var* gene probes, which correspond to the sequences amplified by RT-PCR and cloned from single ring-infected pRBCs 3 and 7. The ethidium-bromide-stained gel is shown on the left. Molecular sizes are given in megabases.

we hybridized transcript sequences from cell 3 (*var-1* and *var-9*) and cell 7 (*var-1*, *var-10* and *var-11*) to PFGE (pulsed field gel electrophoresis)-separated chromosomes of four clonally related but phenotypically distinct FCR3S parasites (rosetting-positive: FCR3S1, FCR3S1.2; rosetting-negative: FCR3S1.6, FCR3S/a) and of 3D7A, an unrelated rosetting-negative parasite (Fig. 4). The *var-1* probe hybridized to chromosome 4, which was of similar size in each of the FCR3S parasites regardless of the rosetting phenotype. The *var-9* sequence from cell 3 hybridized to a chromosome in the compression region and to an additional chromosome of the FCR3S1.6 clone. The other two *var* sequences cloned from cell 7 (*var-10* and *var-11*) hybridized to chromosomes distinct from chromosome 4. None of these probes hybridized to the chromosomes of the parasite 3D7A, indicating that this parasite has a completely different *var* gene repertoire. Our data show that active transcription of multiple *var* genes occurs on different chromosomes at an early stage after invasion of erythrocytes.

As only one *var* transcript was seen in trophozoite-infected red blood cells, we determined whether only one PfEMP-1 type is expressed at the erythrocyte surface. We analysed surface-radiolabelled, infected erythrocytes by SDS/polyacrylamide gel electrophoresis and autoradiography. Only one labelled polypeptide (relative molecular mass ~285,000; M_r 285K) with the characteristics of PfEMP-1 was detected on the surface of red blood cells infected with the FCR3S1.2 clone (Fig. 5), a result consistently found in more than ten independent labelling experiments. To confirm that the polypeptide was PfEMP-1, we performed immunoprecipitations with six monospecific antisera directed either against the semivariable DBL-1 domain of our previously cloned FCR3S1.2 *var-1*, or against the highly conserved acidic terminal sequence of PfEMP-1, which has been found in all *var* genes described so far^{9,10,13,15,16}. We also used three human sera from clinically immune adults from a malaria holoendemic area in Liberia. In all cases, a single polypeptide of M_r ~285K was immunoprecipitated from the labelled FCR3S1.2-infected red blood cells, whereas no other PfEMP-1 precipitated with either of the sera. Thus, although several *var* genes are transcribed in the course of the parasite's intra-erythrocytic development, only one PfEMP-1 type seems to be transported to the host cell surface.

The ability of *P. falciparum* to undergo clonal antigenic variation is well established^{7,17} but, to survive in the human host, *P. falciparum* must maintain a delicate balance between functional conservation

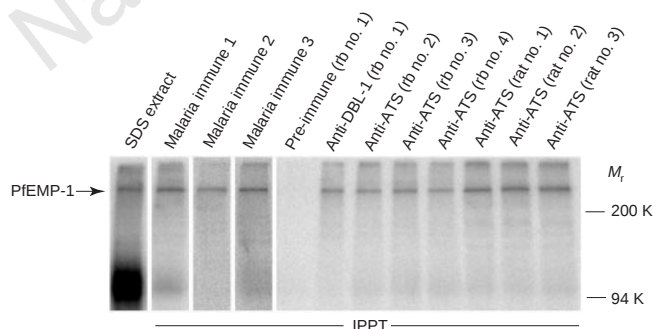


Figure 5 One PfEMP-1 type is expressed on the surface of the trophozoite-infected erythrocyte. Polypeptides from surface-radiolabelled pRBCs infected with mature stages of the FCR3S1.2 parasite were extracted in 2% SDS and immunoprecipitated (IPPT; see Methods). Human sera (malaria immune 1–3) were from adult, clinically immune Liberians with lifelong heavy exposure to malaria (these sera recognize multiple different PfEMP-1 variants in other *P. falciparum* isolates tested; our unpublished data). Monospecific sera to the DBL-1 domain and the highly conserved ATS domain were raised by immunizing rabbits (rb) and rats with fusion proteins encoded by the corresponding fragments of the cloned FCR3S1.2 *var-1* (ref. 13).

(that is, of adhesive phenotypes) and antigenic variation. 'Switching' variants with changed adhesive and antigenic phenotypes, as well as with different PfEMP-1 types, have been generated from the clonal tree of parasites studied here¹⁸. Our RT-PCR data indicate that the switching process is initiated by the activation of several *var* genes at different chromosomal loci early in development, and that it is from this set of activated genes that a new *var* variant is selected as the parasite differentiates into the trophozoite stage. A calculation of the number of *var* transcript species present in the bulk culture of FCR3S1.2 indicates that at least 24 different *var* genes are expressed, out of a total of 45 or more genes estimated to constitute the *var* gene pool of this parasite (see Methods). The data are not as conclusive for the individual ring-stage cells but are still consistent with the simultaneous activation of many *var* genes.

One interpretation of these findings¹¹, is that many, if not all, functional *var* genes are activated in the beginning of the intra-erythrocytic cycle and that any cell has the potential to switch to any *var* gene in the repertoire. However, activation of a set of genes distributed throughout all chromosomes^{14,19} could be a costly and inefficient process, resulting in overflow transcription. This possibility is suggested by the detection of transcripts of the CSP gene, a mosquito-stage-specific antigen, in FCR3S1.2 rings (data not shown) and the transient activation of the sporozoite-specific antigen, STARP, in ring-stage parasites²⁰. Alternatively, large regions of the genome could undergo a process of loose transcription early during intra-erythrocytic development, resulting in a short-lived activation of many, perhaps all, parasite genes. The presence of only the FCR3S1.2 *var-1* transcript in the later, trophozoite, forms shows that there is a preference to switch back to the previously expressed *var* gene (Fig. 6). This bias towards reusing the former PfEMP-1 probably reflects the mechanism of switching and may have a biological role in maintaining an adhesive phenotype, presumably functional in the previous life cycle, as well as in controlling the rate at which new epitopes are exposed to the host. Our results are consistent with the characteristics of clinical immunity to malaria: protection against severe disease is achieved only after a relatively large number of exposures^{21,22}. Taken together, our data show that

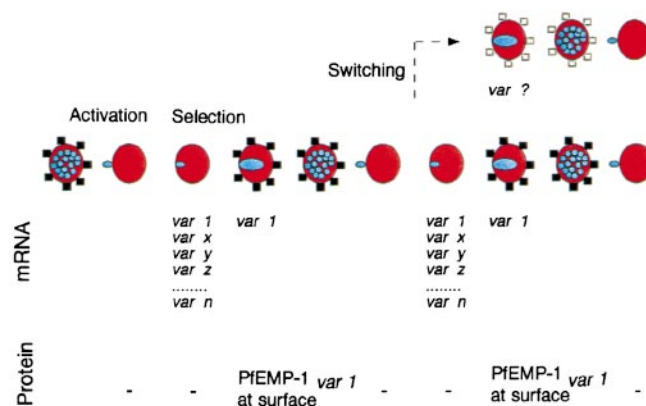


Figure 6 Outline of the activation and selection of *var* genes, and subsequent expression of PfEMP-1 during the intra-erythrocytic phase of asexual *P. falciparum*. Several *var* genes at different chromosomal loci are activated early in development (left). An undefined selection process results in only one *var* transcript being detected as the parasite progresses from the ring to the trophozoite stage, and in the translation, transport and expression at the host cell surface of the corresponding PfEMP-1 polypeptide (black squares). The same *var*/PfEMP-1 is expressed by ensuing parasite generations, although individual parasites may switch the used *var* gene at a frequency of about 10^{-2} per generation, and consequently express a different PfEMP-1 variant (open squares). Red ovals represent red blood cells, blue shapes represent *P. falciparum*.

P. falciparum, in a developmentally regulated manner, activates the transcription of many genes but selects and exposes only one adhesive PfEMP-1 type, a choice that may influence the outcome of the disease. □

Methods

Parasites and isolation of total RNA. The *P. falciparum* parasites were cultured according to standard methods with 10% AB⁺Rh⁺ serum added to the buffered medium (RPMI supplemented with HEPES, gentamycin and sodium bicarbonate). We obtained the clone FCR3S1.2 by micromanipulation from FCR3S1, a parasite previously cloned by limiting dilution²³. The rosetting rate of FCR3S1.2 was routinely 90–95% when used for experiments. FCR3S/a (<5% rosetting) was derived from FCR3S by enrichment of non-rosetting parasites as described¹⁸. Total RNA from FCR3S1.2 and FCR3S/a parasites was purified according to standard procedures from non-synchronized cultures containing predominantly ring-stage parasites.

Collection of single parasites and RT-PCR. We collected individual infected or uninfected erythrocytes from cultures by micromanipulation using 5-μm micropipettes and an inverted light microscope. We carefully examined each collected cell, by repeated grabbing, ejecting and turning, to assess single-infection and maturity stage, or absence of the parasite. Rosette-forming *P. falciparum*-infected red blood cells (pRBCs) were stripped of uninfected cells. Immediately after collection, the cells were frozen on dry ice and subsequently heat-lysed, DNase-treated and amplified by RT-PCR using primers DBL-1.1 and DBL-1.2, as described¹³. Several controls were included in each experiment, including a blank control (without cells), a red-blood-cell control (uninfected erythrocytes), and a control without reverse transcriptase, to rule out the possibility of contamination and amplification due to the presence of genomic DNA. Either in RT-PCR experiments or during direct amplification of genomic DNA, DBL-1.1 and DBL-1.2 primers exclusively amplified *var* DBL-1 fragments, as determined by BLAST analysis of the sequences against the Genebank database. Amplification reactions were analysed in 1% agarose gels.

Amplification of the intron/exon junction region of the *var* 1 gene. As an additional control for eventual genomic DNA amplification, we amplified a 315-bp fragment of FCR3S1.2 *var*-1 complementary DNA, spanning the intron/exon junction of this gene, using the upstream primer P-1 (5'-CTTTCGACTCTACCATCCT-3', located in exon I), the downstream primer L-5.1 (5'-CCAATTCTTCATATTCACCTT-3', located in exon II), and optimized RT-PCR conditions as described¹³.

Southern blot. Chromosomes of different parasites were separated by PFGE and transferred to nylon filters using standard procedures. A 140-bp fragment specific for the FCR3S1.2 *var*-1 was amplified with primers L-2 (5'-TTACATCTCCGTCGCCAG-3') and L-3 (5'-TCGAGCAAGGAGCTTGATA-3'). Hybridization with ³²P-labelled probes (140 bp and 550 bp) was performed with standard methods and washed under stringent conditions (0.1× SSC, 60°C). The filters were analysed by phosphorimager.

Surface labelling of pRBCs and immunoprecipitation. Cultures were radiolabelled with ¹²⁵I by the lactoperoxidase method as described¹⁸. Trophozoite-infected pRBCs were separated by centrifugation in Percoll/Sorbitol gradients and sequentially extracted in 1% TX-100 and 2% SDS. The SDS extract was immunoprecipitated with human, rabbit or rat sera as described²⁴ and monitored by SDS-PAGE and phosphorimager.

Calculation of the number of *var* mRNA species and *var* genes in the population. We used the formula $P = (1 - 1/m)^n$ to estimate the number of unique *var* gene mRNA variants in bulk culture. The probability of finding a unique sequence is represented by P (in this case $P = 8/15 = 0.53$); the number of sampled sequences is represented by n (here $n = 15$); and m represents the number of unique *var* mRNA variants in the population. We assumed that all sequences are expressed at an equal level, that the primers amplify all *var* sequences, and that there is no bias in the PCR under the conditions used. As these assumptions are probably not true (for example, *var*-1 transcripts seem to be over represented), the real number of expressed *var* gene variants is probably higher than the estimated value. We also used the formula described above to obtain a lower limit estimate of the *var* gene pool size of the FCR3S1.2 parasite (with $P = 14/24 = 0.58$ and $n = 24$). We sequenced 24 recombinant clones from genomic FCR3S1.2 DNA amplified with DBL-1.1 and DBL-1.2 primers. All amplified fragments were *var* DBL-1 and 14 sequences were unique.

Received 26 February; accepted 11 June 1998.

1. Miller, L. H., Good, M. F. & Milon, G. Malaria pathogenesis. *Science* **264**, 1878–1883 (1994).
2. MacPherson, G. G., Warrell, M. J., White, N. J., Looareesuwan, S. & Warrell, D. A. Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *Am. J. Pathol.* **119**, 385–401 (1985).
3. Carlson, J. et al. Human cerebral malaria: association with erythrocyte rosetting and lack of anti-rosetting antibodies. *Lancet* **336**, 1457–1460 (1990).
4. Rowe, A., Obeiro, J., Newbold, C. I. & Marsh, K. *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya. *Infect. Immun.* **63**, 2323–2326 (1995).
5. Magowan, C., Wollish, W., Anderson, L. & Leech, J. Cytoadherence by *Plasmodium falciparum*-infected erythrocytes is correlated with the expression of a family of variable proteins on infected erythrocytes. *J. Exp. Med.* **168**, 1307–1320 (1988).
6. Biggs, B. A. et al. Adherence of infected erythrocytes to venular endothelium selected for antigenic variants of *Plasmodium falciparum*. *J. Immunol.* **149**, 2047–2054 (1992).
7. Roberts, D. J. et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357**, 689–692 (1992).
8. Smith, J. D. et al. Switches in expression of *Plasmodium falciparum* *var* genes correlated with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101–110 (1995).
9. Su, X.-Z. et al. The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89–99 (1995).
10. Baruch, D. I. et al. Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of the parasitized human erythrocyte. *Proc. Natl Acad. Sci. USA* **93**, 3497–3502 (1996).
11. Borst, P. et al. Antigenic variation in malaria. *Cell* **82**, 1–4 (1995).
12. Deitsch, K. W., Moxon, E. R. & Welles, T. E. Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol. Mol. Rev.* **61**, 281–293 (1997).
13. Chen, Q. et al. Identification of PfEMP1 as the rosetting ligand of the malaria parasite *Plasmodium falciparum*. *J. Exp. Med.* **187**, 15–23 (1998).
14. Peterson, D. S., Miller, L. H. & Welles, T. E. Isolation of multiple sequences from the *Plasmodium falciparum* genome that encode conserved domains homologous to those in erythrocyte-binding proteins. *Proc. Natl Acad. Sci. USA* **92**, 7100–7104 (1995).
15. Rowe, J. A. et al. *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature* **388**, 292–295 (1997).
16. Hernandez-Rivas, R. et al. Expressed *var* genes are found in *Plasmodium falciparum* subtelomeric regions. *Mol. Cell. Biol.* **17**, 604–611 (1997).
17. Biggs, B. A. et al. Antigenic variation in *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **88**, 9171–9174 (1991).
18. Fernandez, V., Treutiger, C. J., Nash, G. B. & Wahlgren, M. Multiple adhesive phenotypes linked to rosetting-binding of erythrocytes in *Plasmodium falciparum* malaria. *Infect. Immun.* **66**, 2969–2975 (1998).
19. Rubio, J. P., Thompson, J. K. & Cowman, A. F. The *var* genes of *Plasmodium falciparum* are located in the subtelomeric region of most chromosomes. *EMBO J.* **15**, 4069–4077 (1996).
20. Fidock, D. A. et al. Cloning and characterization of a novel *Plasmodium falciparum* sporozoite surface antigen, STARP. *Mol. Biochem. Parasitol.* **64**, 219–232 (1994).
21. Marsh, K. & Howard, R. J. Antigens induced on erythrocytes by *P. falciparum*: expression of diverse and conserved determinants. *Science* **231**, 150–153 (1986).
22. Gupta, S. et al. Parasite virulence and disease patterns in *Plasmodium falciparum* malaria. *Proc. Natl Acad. Sci. USA* **91**, 3715–3719 (1994).
23. Udomsangpetch, R. et al. *Plasmodium falciparum*-infected erythrocytes form spontaneous erythrocyte rosettes. *J. Exp. Med.* **169**, 1835–1840 (1989).
24. van Schravendijk, M. R. et al. Characterization and localization of *Plasmodium falciparum* surface antigens on infected erythrocytes from West African patients. *Blood* **78**, 226–236 (1991).

Acknowledgements. We thank K. Berzins for the gift of malaria immune sera. These studies were supported by grants from the Swedish Medical Research Council and the Karolinska Institutet.

Correspondence and requests for materials should be addressed to M.W. (e-mail: mats.wahlgren@smi.ki.se). Sequences have been deposited in the Genebank under accession numbers AF039275–AF039283 and AF061155.

Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor

Thomas P. J. Garrett*, Neil M. McKern†, Meizhen Lou*, Maurice J. Frenkel†, John D. Bentley†, George O. Lovrecz†, Thomas C. Elleman†, Leah J. Cosgrove† & Colin W. Ward†

* Biomolecular Research Institute and †CSIRO Division of Molecular Science, 343 Royal Parade, Parkville, Victoria 3052, Australia

The type-1 insulin-like growth-factor receptor (IGF-1R) and insulin receptor (IR) are closely related members of the tyrosine-kinase receptor superfamily¹. IR is essential for glucose homeostasis², whereas IGF-1R is involved in both normal growth and development and malignant transformation³. Homologues of these receptors are found in animals as simple as cnidarians⁴.

† Present address: CRC for Tissue Growth and Repair, Gouger Street, Adelaide, Australia.

The epidermal growth-factor receptor (EGFR) family is closely related to the IR family and has significant sequence identity to the extracellular portion we describe here. We now present the structure of the first three domains of IGF-1R (L1–Cys-rich–L2) determined to 2.6 Å resolution. The L domains each consist of a single-stranded right-handed β -helix. The Cys-rich region is composed of eight disulphide-bonded modules, seven of which form a rod-shaped domain with modules associated in an unusual manner. The three domains surround a central space of sufficient size to accommodate a ligand molecule. Although the fragment (residues 1–462) does not bind ligand, many of the determinants responsible for hormone binding and ligand specificity map to this central site. This structure therefore shows how the IR subfamily might interact with their ligands.

The ligands IGF-I, IGF-II and insulin share a common architecture and competitively crossreact with IGF-1R and IR³. It is therefore likely that they bind to the receptors in a structurally equivalent fashion. Ligand binding has previously been demonstrated only for mature receptors, receptor ectodomains and monomeric α -chain⁶. Furthermore, residues 704–715 of IR have been strongly implicated in ligand binding⁷, so it is perhaps not surprising that the fragment crystallized⁸ does not bind IGF-I.

The fragment adopts an extended bilobal structure ($40 \times 48 \times 105$ Å) with the L domains at either end (Fig. 1). The Cys-rich region runs two-thirds of the length of the molecule, making contact along

the length of L1 but having very little contact with L2. The L domains (residues 1–150 and 300–460, each approximately $24 \times 32 \times 37$ Å) consist of a single-stranded right-handed β -helix, capped on the ends by short helices and disulphide bonds. Each domain is shaped like a loaf of bread, with the base formed from a flat six-stranded β -sheet five residues long (green in Figs 1 and 2a), the sides formed from β -sheets three residues long (blue and yellow in Figs 1 and 2a), and the top being irregular. A conserved loop of 7–8 residues is inserted between the fourth 'blue' and 'green' strands. The right-angle bends between the sheets appear to be dictated by sequence (Fig. 2a). On each 'yellow' sheet there is a short asparagine ladder (for example, residues 395, 411 and 452 in L2) with the side chains forming hydrogen bonds back to main-chain atoms of the green strands to create the bend. Between the blue and green β -strands, conserved glycines appear to be structurally important as mutation of these residues compromises folding of the receptor⁹. The two domains are superposable with a root-mean-squared deviation (r.m.s.d.) in position of 1.6 Å for 109 C α atoms (Fig. 3). A sequence motif, corresponding to part of the β -helical repeat, was identified previously¹⁰. Some green β -strands were correctly predicted but the ' α -helix-like' motif preceding the conserved glycine corresponds to a pair of bends and a blue β -strand.

The L domains differ at their carboxy-terminal ends. For L1, the indole ring of Trp 176 from the Cys-rich region is inserted between the last two turns of β -helix into the hydrophobic core of the domain and the C-terminal helix of L1 becomes vestigial (Fig. 3). The sequence motif for residues forming the Trp pocket in L1 does not occur in L2 of the IR family (red in Fig. 2a). Furthermore, in EGFR, which has an additional Cys-rich region after the L2 domain^{10,11}, the motif can be found in both L domains and the Trp is conserved in both Cys-rich regions (Fig. 2b).

Comparison of the L domains with other protein structures¹² identified two folds, each of which has a right-handed superhelical topology: horseshoe-shaped ribonuclease inhibitor¹³ and β -helix structures such as pectate lyase¹⁴. These domains are substantially larger than L domains but all have similarly sized superhelical repeats (22–25 residues). Each has a regularly stacked hydrophobic core composed primarily of aliphatic side chains and a version of an asparagine ladder¹³. It might be expected that β -helix domains are more closely related to L domains, as they have three β -sheets and are often capped by amino-terminal α -helices, but this is not the case. The cross-sections of the β -helix domains are triangular or V-shaped, whereas L domains and ribonuclease inhibitor have rectangular cross-sections. Furthermore, the large β -sheet of L domains (green in Figs 1 and 2a) and the sheet of ribonuclease inhibitor are untwisted, whereas sheets in pectate lyase, for example, have a significant twist. Thus the fold of the L domains appears marginally more similar to ribonuclease inhibitor than a β -helix but constitutes a new type of right-handed superhelical structure.

As anticipated¹¹, the Cys-rich domain is composed of modules (defined in Fig. 2b) with disulphide-bond connectivities resembling parts of the tumour-necrosis-factor receptor and laminin. The first module sits at the end of L1, the remaining seven forming a curved rod running diagonally across L1 and reaching to L2 (Fig. 1). The strands in modules 2–7 run roughly perpendicular to the axis of the rod, rather like laminin¹⁵, but the modular arrangement of the Cys-rich domain is different from other Cys-rich proteins for which structures are known. Modules 1–3 of IGF-1R have a common core, containing a pair of disulphide bonds (Cys 1–3 and 2–4), but show considerable variation in the loops (Fig. 2b). Modules 4–7 have a β -finger motif, each composed of three polypeptide strands, with the first and third strands disulphide bonded and the second and third forming a β -ribbon. The four β -ribbons line up antiparallel to form a tightly twisted eight-stranded β -sheet (Figs 1 and 4). Module 6 deviates from the common pattern, with the first segment being replaced by an α -helix followed by a large loop that is implicated in

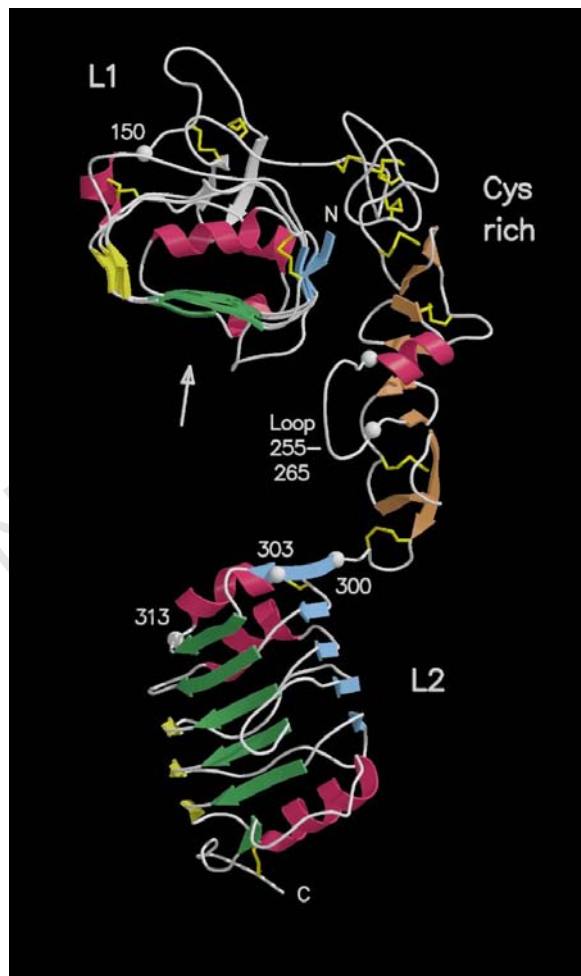


Figure 1 Polypeptide fold for residues 1–459 of IGF-1R. The L1 domain is at the top, viewed from the N-terminal end. The space at the centre is of sufficient size to accommodate IGF-I. Helices are indicated by curled ribbon and β -strands by broad arrows. Residues referred to in the text are labelled and marked as white spheres, and disulphide bonds are shown as yellow sticks. The white arrow near the centre of the figure indicates the view direction for Fig. 5.

ligand binding (see below). The final module is a disulphide-linked bend of five residues. The first two types of modules have previously been described as parts of an EGF module, but here each type on its own forms a regular, extended structure. These modules must therefore be considered to be the minimal building blocks of Cys-rich domains, and proteins such as laminin¹⁵ can be considered as a series of repeat units, each composed of a small number of modules.

Attempts have been made to locate the binding site of IGF-I (and insulin) by examining natural^{9,16} and site-directed mutants^{7,17}, chimaeric receptors^{18–20} and by chemical crosslinking studies²¹. IGF-1R/IR chimaeras not only show which regions of the receptors account for ligand specificity, but also provide an efficient means of identifying some parts of the hormone-binding site. Paradoxically, replacing residues 1–62 of IGF-1R with 1–68 of IR confers insulin-

binding ability on IGF-1R¹⁸, whereas substitution of IR 198–300 in the Cys-rich region with IGF-1R 191–290 allows this receptor to bind IGF-I¹⁹. Thus a receptor can be constructed that binds both IGF-I and insulin with comparable affinity. From the structure, it is clear that a ligand bound in the central space could contact both these regions.

The specificity determinant in the Cys-rich region can be further limited²⁰ to residues 223–274, and this region (modules 4–6) includes a large, mobile loop (for C α atoms of residues 255–265; $B = 63 \text{ \AA}^2$) extending into the central space (Fig. 1). This loop is four residues larger in IR and is stabilized by an additional disulphide bond. In IR, this loop may serve to perturb IGF-I, reducing its affinity but still allowing the smaller insulin molecule to bind tightly. Complementary changes are also observed in the

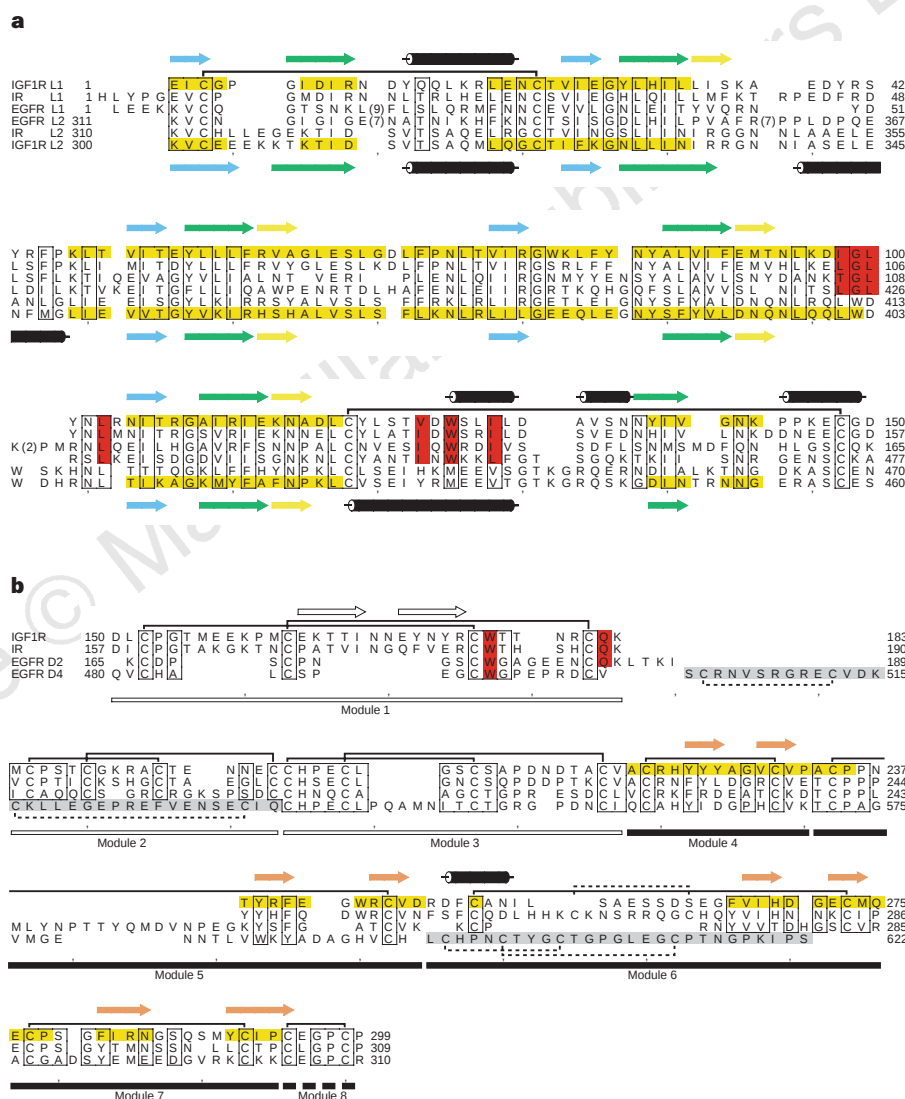


Figure 2 Amino-acid sequences of IGF-1R and related proteins. **a**, The L1 and L2 domains of human IGF-1R were aligned using structural considerations, and residues shaded yellow are structurally equivalent. IR and EGFR are shown based on a sequence alignment to IGF-1R. Numbers in parentheses indicate where amino acids have been omitted. Positions with conserved physicochemical properties of amino acids are boxed, and residues that form the Trp 176 pocket are in red. Secondary structure elements for L1 (above the sequences) and L2 (below) are indicated as cylinders for helices and arrows for β -strands. Strands are colour coded according to the β -sheet to which they belong, as shown in Fig. 1. Disulphide bonds are also indicated. **b**, Cys-rich domains of human IGF-1R, IR and EGFR (domains 2 and 4) are aligned using sequence and

structural considerations. Secondary structural elements and disulphide bonds for IGF-1R are drawn above the sequences. Dashed bonds indicate disulphide links present only in IR (above) or EGFR (below). Modules have been defined as minimal discretely folded subunits and the three types, based on structural similarity, are labelled below the sequences as open, filled and broken bars. Boxed residues show conservation of physicochemical properties. Structurally conserved residues for modules 4–7 are shaded yellow with cysteine residues being equivalent in the first and third strands and residues 225, 240, 268 and 283 in the second strand. Residues from EGFR that do not conform to the pattern are shaded grey, and the conserved Trp 176 and the semi-conserved Gln 182 nearby are shaded red. This figure was prepared using ALSCRIPT.

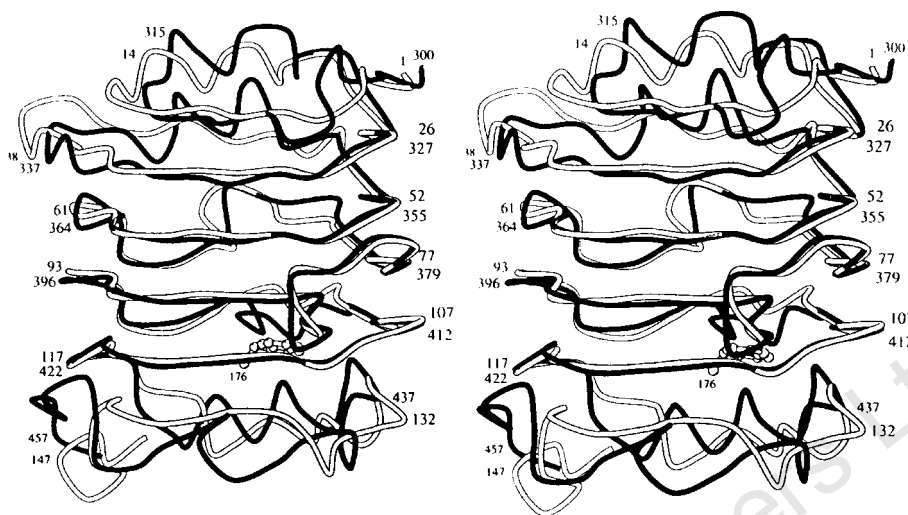


Figure 3 Stereo view of a superposition of the L1 (white) and L2 (black) domains. Residue numbers above are for L1 and below for L2. The side chain of Trp 176, which protrudes into the core of L1, is shown by a ball-and-stick representation.

ligand, IGF-I, in which replacement of the C region (residues 28–37) by a four-glycine linker reduced affinity for IGF-1R 40-fold but increased affinity for IR 5-fold²².

Alanine substitutions in IR identified four regions of L1 that are important for insulin binding¹⁷, of which three occur on successive turns of the β -helix and the fourth lies on the large conserved bulge on the large β -sheet. These regions create a footprint on the first half of the large (green) β -sheet of the L1 domain where insulin binds (Fig. 5); this footprint faces into the central space (Fig. 1). Residues further down the sheet, which are conserved in IGF-1R, could also be important but as yet have not been examined. The L2 domain presents the N-terminal end of its β -helix to the central space. On this face, the only mutation observed so far is S323L of IR (S313 of IGF-1R), which gives rise to Rabson–Mendenhall syndrome and severe insulin resistance²³. This mutant affects only insulin binding, and structurally it lies in the middle of a region that is quite well conserved in both receptors (residues 309–318 and 332–345 of IGF-1R). Therefore, although this region may be part of the hormone-binding site, it would not have been detected by chimaeras.

The distance between the putative hormone-binding regions on L1 and L2 is approximately 30 Å (Fig. 1), so the domains appear a little too far apart to bind IGF-I or insulin. Furthermore, as L2 makes no significant contact with the L1 and Cys-rich domains in the crystal, in solution it probably has no preferred orientation with respect to the rest of the fragment. The position of L2 observed here is dictated by crystal packing contacts. For example, there is a deep cleft between part of the Cys-rich domain (residue 262) and L2 (residue 303) that is occupied by a loop from a neighbouring molecule. Possibly, in the whole receptor there is, a substantial

rotation of L2 about residue 300 to allow the hormone-binding regions of L1 and L2 to contribute to different hormone-binding sites. However, all the residues in this fragment that are known directly to affect hormone binding map to the three faces of the molecule that point into the central space. A rotation of L2 by 25° would be sufficient to close the cleft between residues 262 and 303 and allow a hormone such as IGF-I to contact all three regions.

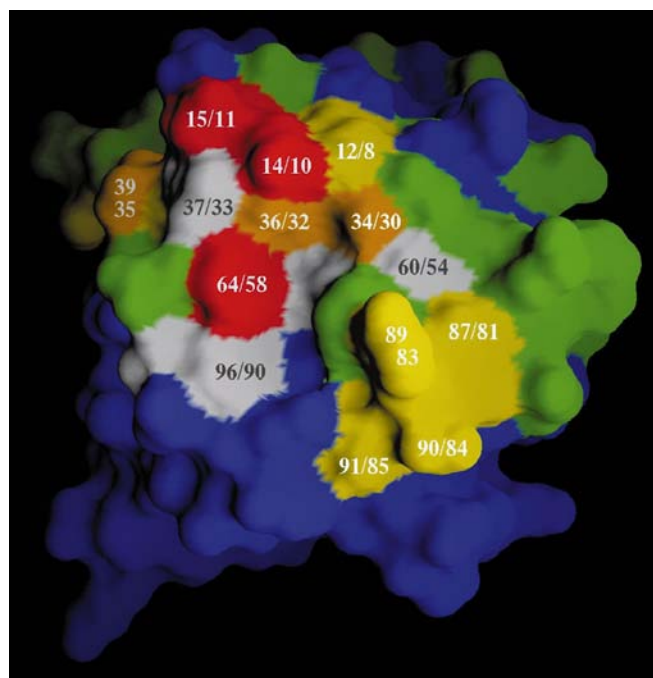


Figure 5 Surface diagram of the L1 domain of IGF-1R showing the hormone-binding footprint. The view direction for this figure is indicated in Fig. 1, and the domain is shown in a similar orientation to Fig. 3. The mutation²⁰ L87A and four regions (residues 12–15, 34–44, 64–67 and 89–91 or IR), previously shown to be important in insulin binding to IR¹⁷, correspond to a patch of residues on the large β -sheet. Residue numbers for IR/IGF-1R are given and residues are coloured according to the magnitude of $K_a(\text{mutant})/K_a(\text{wild type})$: red, >40; orange, 10–40; yellow, 2.4–10; green, <2.5; white, non-secreting; blue, untested. All mutants on the opposite face of the domain do not affect insulin affinity.

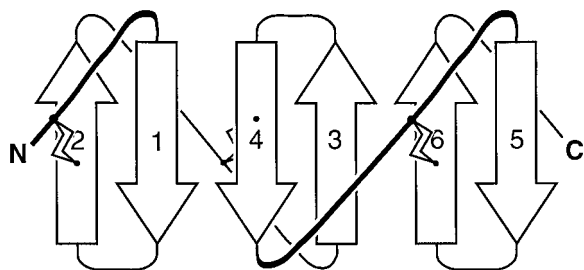


Figure 4 Diagram of the association of β -finger motifs. The β -strands are drawn as arrows and disulphide bonds as zig-zags.

Table 1 Summary of crystallographic data

Data set	Resolution (Å)	Mean I/σ	R _{merge} *	Completeness (multiplicity)	No. of sites	R _{cullis} †	Phasing power‡	f.o.m.§
Native	2.6	18.7	0.064	0.996 (4.1)				0.47/0.71
PIP	3.0	15.8	0.060	0.982 (2.2)	3	0.66	1.71	
UO ₂ Ac ₂	4.5	7.5	0.095	0.989 (2.3)	2	0.82	1.17	
Refinement resolution (Å)	No. of reflections (free)		No. of atoms		R _{cryst}	R _{free}	Bonds¶ (Å)	Angles¶ (Å)
7.0–2.6	24,270 (2,693)		3,903		0.237	0.304	0.017	0.048

PIP, di-μ-iodobis(ethylenediamine)diplatinum dinitrate; UO₂Ac₂, uranyl acetate.

*R_{merge} = $\sum_h \sum_j |I_{hj} - \bar{I}_h| / \sum_h \sum_j I_{hj}$, where I_{hj} is an intensity measurement j and $\bar{I}_h$ is the mean intensity for reflection h .

†R_{cullis} = $\sum_h |F_{PH} - F_P| - |F_{Hcalc} - F_H| / \sum_h |F_{PH}| + |F_P|$, where F_{PH} , F_P and F_{Hcalc} are, respectively, derivative, native and heavy-atom structure factors for centric reflections h .

‡Phasing power = $\sum_h |F_{Hcalc}| / \sum_h |F_P|$, where F_{Hcalc} is defined above and ϵ is the lack of closure.

§f.o.m. (figure of merit) = $(\cos(\Delta\alpha_h))$, where $\Delta\alpha_h$ is the error in the phase angle for reflection h . Values are given before and after density modification at 3.0 and 2.8 Å resolution, respectively.

||R_{cryst} and R_{free} are defined in ref. 28.

¶R.m.s. deviation from ideal bond and angle-related (1–3) distances.

Unlike most other cell-surface receptors, which dimerize upon binding ligand, members of the IR subfamily pre-exist as disulphide-linked dimers. Thus a small domain movement such as this may indeed be necessary to transmit the ligand-binding signal into the cell.

The structure of the L1–Cys-rich–L2 domains of IGF-1R provides the first information (to our knowledge) about the extracellular portion of a member of the insulin receptor family. Each domain displays new architectural features that are common both to this family and the EGF receptor family. Although this polypeptide encompasses only the first half of the IGF-1R ectodomain, it contains the major specificity determinants for its ligands. Data from chimaeras and single-site mutants indicate that the hormone probably binds at the centre of this fragment, making contact with all three domains. Even though the picture is not complete, this structure gives a first glimpse of how IGF-I, IGF-II and insulin might interact with their cell-surface receptors to produce their metabolic and mitogenic effects. □

Methods

Crystals⁸ (space group, $P2_12_12_1$, $a = 77.39$ Å, $b = 99.72$ Å, $c = 120.29$ Å) were cryo-cooled to -170°C in a mother liquor containing 20% glycerol, 2.2 M ammonium sulphate and 100 mM HEPES, pH 7.5. Diffraction data were recorded on Rigaku RAXIS IIC or IV area detectors using a Siemens M18XHF X-ray generator with Yale/MSC mirror optics and reduced using DENZO²⁴. Diffraction was notably anisotropic for all crystals examined. Phasing by multiple isomorphous replacement (MIR) was performed with PROTEIN²⁵ using anomalous scattering for both UO₂ and PIP derivatives. Statistics are given in Table 1. In the initial MIR map, regions of protein and solvent could clearly be seen but the path of the polypeptide was not obvious. After phase extension to 2.8 Å resolution by solvent flattening and histogram matching²⁶, the electron density map was considerably better than the MIR map, probably because of the large solvent content of the crystals (66%). The structure was traced with O²⁷ and rebuilding was alternated with rounds of energy minimization using X-PLOR 3.851 (ref. 28). In the third round of refinement an overall anisotropic temperature factor was refined, the magnitudes of the semi-axes (8.1, 1.0 and -13.1 Å²) reflecting the substantial anisotropy of diffraction observed earlier. Later in the refinement, ten residues from the *c-myc* tail were located and large peaks in a difference map ($>4\sigma$) were included as ordered solvent where stereochemically reasonable. The final cycles of refinement were performed with REFMAC²⁹, giving $R_{\text{cryst}} = 0.237$ and $R_{\text{free}} = 0.304$. The structure contains 469 amino acids, 10 carbohydrate residues (at residues 21, 105, 214 and 284), 8 sulphate ions and 48 water molecules. For residues with $B(\text{C}\alpha) > 75$ Å² atomic positions are less reliable (36–42, 154–160, 258–260, 305–306, 336–342, 404–407 and 457–459). For 24 residues, including residue 83, there was insufficient electron density to include the whole side chains. The electron density is weak for residues 457–459 and residues 460–468 appear disordered.

Received 4 March; accepted 21 May 1998.

- Ullrich, A. *et al.* Insulin-like growth factor 1 receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. *EMBO J.* **5**, 2503–2512 (1986).

- De Meyts, P. The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling. *Diabetologia* **37**, 135–148 (1994).
- Baserga, R. Controlling IGF-receptor function: a possible strategy for tumor therapy. *Trends Biotechnol.* **14**, 150–152 (1996).
- Steele, R. E., Lieu, P., Mai, N. H., Shenk, M. A. & Sarraf, M. P. Jr Response to insulin and the expression pattern of a gene encoding an insulin receptor homologue suggest a role for an insulin-like molecule in regulating growth and patterning in *Hydra*. *Dev. Genes. Evol.* **206**, 247–259 (1996).
- Torres, A. M. *et al.* Solution structure of human insulin-like growth factor II—relationship to receptor and binding protein interactions. *J. Mol. Biol.* **248**, 385–401 (1995).
- Schaefer, E. M., Siddle, K. & Ellis, L. Deletion analysis of the human insulin receptor ectodomain reveals independently folded soluble subdomains and insulin binding by a monomeric α -subunit. *J. Biol. Chem.* **265**, 13248–13253 (1990).
- Mynarcik, D. C., Yu, G. Q. & Whittaker, J. Alanine-scanning mutagenesis of a C-terminal ligand binding domain of the insulin receptor alpha subunit. *J. Biol. Chem.* **271**, 2439–2442 (1996).
- McKern, N. M. *et al.* Crystallization of the first three domains of insulin-like growth factor I receptor. *Protein Sci.* **6**, 2663–2666 (1997).
- Taylor, S. I. *et al.* Mutations in the insulin receptor gene: update 1994. *Endocr. Rev.* **2**, 58–65 (1994).
- Bajaj, M., Waterfield, M. D., Schlessinger, J., Taylor, W. R. & Blundell, T. On the tertiary structure of the extracellular domains of the epidermal growth factor and insulin receptors. *Biochim. Biophys. Acta* **916**, 220–226 (1987).
- Ward, C. W., Hoyne, P. A. & Flegg, R. H. Insulin and epidermal growth factor receptors contain the cysteine repeat motif found in the tumour necrosis factor receptor. *Proteins Struct. Funct. Genet.* **22**, 141–153 (1995).
- Holm, L. & Sander, S. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
- Kobe, B. & Deisenhofer, J. Proteins with leucine-rich repeats. *Curr. Opin. Struct. Biol.* **5**, 409–416 (1995).
- Yoder, M. D., Lietzke, S. E. & Jurnak, F. Unusual structural features in the parallel β -helix in pectate lyases. *Structure* **1**, 241–251 (1993).
- Stetefeld, J., Mayer, U., Timpl, R. & Huber, R. Crystal structure of three consecutive laminin-type epidermal growth factor-like (LE) modules of laminin gamma-1 chain harboring the nidogen binding site. *J. Mol. Biol.* **257**, 644–657 (1996).
- Taylor, S. I. Lilly lecture: molecular mechanisms of insulin resistance. Lessons from patients with mutations in the insulin receptor gene. *Diabetes* **41**, 1473–1490 (1992).
- Williams, P. F., Mynarcik, D. C., Yu, G. Q. & Whittaker, J. Mapping of an NH2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis. *J. Biol. Chem.* **270**, 3012–3016 (1995).
- Kjeldsen, T. *et al.* Ligand specificities of insulin receptor and the insulin-like growth factor I receptor reside in different regions of a common binding site. *Proc. Natl Acad. Sci. USA* **88**, 4404–4408 (1991).
- Schäffer, L. *et al.* Interactions of a hybrid insulin-like growth factor-I analog with chimeric insulin/Type I insulin-like growth factor receptors. *J. Biol. Chem.* **268**, 3044–3047 (1993).
- Gustafson, T. A. & Rutter, W. J. The cysteine-rich domains of the insulin and insulin-like growth factor I receptors are primary determinants of hormone binding specificity. *J. Biol. Chem.* **265**, 18663–18667 (1990).
- Kurose, T. *et al.* Cross-linking of a B25 azidophenylalanine insulin derivative to the carboxy-terminal region of the alpha-subunit of the insulin receptor. Identification of a new insulin-binding domain in the insulin receptor. *J. Biol. Chem.* **269**, 29190–29197 (1994).
- Bayne, M. L. *et al.* The C-region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type I IGF receptor. *J. Biol. Chem.* **264**, 11004–11008 (1989).
- Roach, P. *et al.* A novel human insulin receptor gene mutation uniquely inhibits insulin binding without impairing posttranslational processing. *Diabetes* **43**, 1096–1102 (1994).
- Otwinski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1996).
- Steigeman, W. Dissertation, Technical Univ. Munich (1974).
- Cowan, K. DM: an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM newsletter. *Protein Crystallogr.* **31**, 34–38 (1994).
- Jones, T. A., Cowan, S., Zou, J.-Y. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brunger, A. T. *X-PLOR reference manual 3.851* (Yale Univ., New Haven, CT, 1996).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
- Nakae, J., Morioka, H., Ohtsuka, E. & Fujieda, K. Replacements of leucine 87 in human insulin receptor alter affinity for insulin. *J. Biol. Chem.* **270**, 22017–22022 (1995).

Acknowledgements. Financial support was provided under the Generic Technology component of the Industry Research and Development Act 1986 and from Biota Diabetes Research Pty. We thank P. M. Colman for support and discussions.

Correspondence and requests for materials should be directed to C.W.W. (e-mail: colin.ward@molsci.csiro.au) or T.P.J.G. (e-mail: tom.barrett@bioresci.com.au). Coordinates will be deposited with the Protein Data Bank, Brookhaven.

Instrumental advances

This selection of new instrumentation for the laboratory majors on spectroscopy and includes an FT-IR probe designed for inhospitable environments and deuterium lamps achieving notably low noise levels.

SSB-compatible instruments

From Keithley

Universal serial-bus (USB), data communication for highly accurate data acquisition and precision measurement

With its sixteen SmartLink miniaturized measurement instruments, this is the first family of instruments offered by Keithley that is capable of connections to USB-equipped PCs and networks. Compatible devices can be connected and disconnected under full power if necessary for easy maintenance. The USB standard provides constant data transmission rates up to 12 megabytes per second, allowing data collection in demanding applications that require short sample intervals. Such applications include those that entail high-fidelity audio and highly compressed video or high-frequency shock and vibration signals, as well as transient waveforms. The SmartLink drivers for the personal computer enable plug-and-play without requiring reconfiguration of the PC or the installation of a communications card or other interface. When the appropriate driver is loaded into a PC with a USB interface, the computer will recognize the USB-equipped instrument as soon as it is plugged in. Up to 127 instruments can be daisy-chained in a single USB network.

Reader Enquiry No. 100

Spectrometer verification

From Hewlett Packard

Chemical and accessory kits that are designed to provide cost-effective, operational and performance verification of ultraviolet/visible spectrometers

These chemical kits are aimed at pharmaceutical analysts needing to conform closely with quality and regulatory requirements when performing ultraviolet (UV)/visible measurements. They may be used with any spectrophotometer and comprise pre-prepared solutions in sealed ampoules. The solutions are specified by the US Pharmacopoeia and European Pharmacopoeia, and include holmium oxide for wavelength accuracy measurement, potassium dichromate for photometric accuracy measurement, sodium nitrite, sodium iodide and potassium chloride for stray-light measurements at 340, 220 and 200 nm, and toluene in hexane for resolution measurement. Each standard includes a certificate of analysis for traceability. The hardware accessory kit is designed specifically for use with the HP 8453 UV/visible spectrophotometers and comprises flow cells, tubing and syringes, as well as an adjust-



Pharmacopoeia-friendly: HP's verification kits.

ment tool for the multi-cell transport accessory and an instruction manual.

Reader Enquiry No. 101

Spectramax 190

From Molecular Devices

Designed to be a high-performance microplate spectrometer

This spectrometer uses the PathCheck sensor to measure the actual path length of each sample in the microplate and normalizes the absorbance values to a 1-cm cuvette. This feature can be used to check the performance of pipetting devices, determine the concentration of the sample directly, expand the dynamic range of the assay to greater than six absorbance units and detect pipetting errors to evaluate outlier data. The instrument operates over the wavelength range 190–850 nm with a bandwidth of 5 nm, giving the multi-sample capability of a microplate reader with the tunability of a spectrophotometer.

Reader Enquiry No. 102

Vision 32

From Unicam

Operating software that assists with UV-visible validation

Controlling Varian's UV300 and UV 500 spectrometers is a 32-bit package, developed for operation under Windows95 or NT. This package was designed around the company's 'validation qualification calibration' concept, which allows data to be automatically timed and date stamped, and where any changes are recorded in an audit trail. An integral report composer places and sizes methods, graphs, tables and data in the desired position, and the 'uvcalc' spread-

sheet performs automatic post-run calculations within or between different batches.

Reader Enquiry No. 103

GCQ plus

From ThermoQuest Finnigan

A benchtop ion-trap mass spectrometer featuring external ionization technology

GCQplus features MS/MS capability as standard, allowing positive identification of compounds in complex matrices and minimizing sample preparation steps. The instrument incorporates CE Instruments' new Trace CG 2000 for high chromatographic performance. There is also Windows NT software that can be customized for specific applications. A direct probe system is available for quick screening of samples or applications that do not require a GC column.

Reader Enquiry No. 104

SRS 3400

From Bruker AXS

X-ray fluorescence and XRF analysis featuring Spectraplus 32-bit software

This system is designed for straightforward multi-element analysis from beryllium to uranium. The SRS 3400 X-ray spectrometer uses stable 32-bit software, has 4-kW, 150-mA excitation and includes matrix corrections based on the fundamental parameter method — all precalibrated for standardless XRF analysis. The system is said to ensure non-destructive, environmentally safe analysis, and does not require the dissolution of samples or the disposal of the hazardous waste solvents needed for 'wet chemistry'. Sample preparation is simple, whether for analysis of metals, pressed pellets, fused bead or liquids.

Reader Enquiry No. 105

FT-IR probe

From Nicolet Instruments

A fibre-optic probe for analysing samples that do not fit into a sample compartment

Also suited to samples that do not conform to traditional sampling accessories, this



It's so rugged: Nicolet's new probe goes anywhere.

rugged FT-IR probe offers flexibility for analyses involving hazardous environments, samples that produce toxic vapours, reaction monitoring, restricted access sampling vessels, non-destructive sampling and oddly shaped samples. The probe, cables and optical interface are optimized to give double the throughput and sensitivity of earlier versions. The company also offers mid-IR, fibre-optic sampling probes for the analysis of liquids (aqueous and non-aqueous), sheet materials, reflective solids and micro-samples. Probes that use ATR or specular reflection techniques are also available.

Reader Enquiry No. 106

Dart Mach3Xe

From Mass Spectrometry

A mass spectrometry data system with data acquisition in real time (DART)

Based on the latest RISC Sun workstation, this integrated system is designed for low-cost data acquisition, processing and system control. It is compatible for a wide range of mass spectrometers, including all Kratos instruments. This system is said to extend the lifetime and productivity of a mass spectrometer. Moreover, it can replace existing computers and interfacing electronics as they become unreliable and costly to maintain. It is also said to add extra capability for data acquisition and processing. The compact casing is designed for benchtop PC or floor-standing, midi-PC tower, allowing effective use of limited laboratory space. Based on the open Windows desktop environment, run under UNIX, this software provides high-speed graphics in a multiprocessing, 'windowing' system. Data acquisition and instrument control are tuned to provide stability, accuracy, excellent peak definition (irrespective of scan speed) and detection of the smallest and largest components of a sample in a single analysis.

Reader Enquiry No. 107



VAST improvement: 96-well plates meet NMR.

VAST for NMR

From Varian

An NMR automation technique for high-volume applications

VAST, or 'versatile automated sample transport' combines automation hardware and software in one system allowing researchers to quickly and easily introduce liquids directly from the sample handler to the NMR magnet, without the need for dedicated tubes or expensive deuterated solvents. The system brings the 96-well microtitre plate format to NMR through the use of the Microflow NMR probe. It is well suited to high-throughput screening and analysis, such as the construction of combinatorial chemistry libraries and NMR analysis of structure-activity relationships. A typical run is stated to produce up to 500 spectra in a day. The system also features a Gilson 215 liquid handler, which is controlled by the spectrometer and allows full access to the automation capabilities of Varian's VNMR software. The Gilson x-y-z robot holds up to 12 microtitre plates at one time.

Reader Enquiry No. 108

DS200 series lamps

From Heraeus Noblelight

Deuterium lamps that were created for high stability and improved noise levels

These lamps are stated to offer tenfold lower noise levels than earlier lamps, and should provide high stability and low noise levels for

high-resolution HPLC detectors. The series has stated noise levels of 2.5×10^{-5} AU and measured levels of 5×10^{-6} AU after running in a UV detector for 1,100 hours. Ozone noise has been eliminated through the use of high transmission quartz glass material and by limiting the spectral output to 200 nm, using a built-in filter. A low noise level is achieved for all aperture sizes, says the manufacturer, as well as with 'shine-through' lamps. The lamps can be used in any installation in which deuterium lamps of the previous series are used and cover spectral outputs from 190 to 400 nm, with aperture sizes of 1.0 to 400 nm.

Reader Enquiry No. 109

PlasmaQuad ICP-MS

From VG Elemental

Enhanced geological analysis using an ultraviolet microprobe laser ablation ICP-MS

PlasmaQuad ICP-MS is a sensitive technique for difficult applications such as the analysis of fluid inclusions and transparent minerals. The UV microprobe is said to offer high-quality optics and excellent spatial resolution, as well as active focusing through the use of a high-precision stage. Time-resolved analysis software for fast data acquisition, as well as the S-Option interface, which enhances sensitivity for laser ablation ICP/MS by a stated factor of 5-10 times, complements this. ICP-MS has been used to assist multi-element analysis of single fluid inclusions, used as a geological record of the fluid content of the Earth's crust and to aid modelling of ore deposit structures. The UV laser and software combination is said to allow accurate targeting of a fluid inclusion beneath the sample surface, achieving multi-element analysis of the fluid in the range of p.p.b. to p.p.m. Using the microprobe, data can be gathered on difficult mineral types, such as light-coloured apatite, where the rare-earth element content is often high.

Reader Enquiry No. 110

ADVERTISEMENTS

Check out our homepage at <http://www.resgen.com>

BACs
YACs
PACs
cDNAs

Research Genetics provides clones and colony membranes for a number of human, mouse, and rat libraries as well as a custom screening service for the libraries.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

New Severn Super Sequencer Acrylamide Gel Solution

Severn Super Sequencer is a new improved gel solution for use in automated DNA sequencing. Complete with denaturant and buffer, it comes in a series of 'ready to use' (RTU) formulations. Simply pour and Polymerise.

So in the race to sequence, don't walk, RUN!

With Severn Super Sequencer - Available in 100ml, 500ml and 5000ml volumes.

Trial 100ml available on request.

Severn Biotech Ltd

Severn Biotech Ltd, Unit 2, Park Lane, Stourport Road, Kidderminster, Worcs DY11 6TJ, U.K.
Tel: (01562) 825286 Fax: (01562) 825284 E-Mail: 101626.2102@Compuserve.com

UK Manufacturers of Liquid Acrylamides, Buffers and Electrophoretic Media